



HAL
open science

Étude du processus pathogénique associé à la maladie de Chagas à l'aide d'analyses multi-omiques

Pauline Brochet

► **To cite this version:**

Pauline Brochet. Étude du processus pathogénique associé à la maladie de Chagas à l'aide d'analyses multi-omiques. Bio-Informatique, Biologie Systémique [q-bio.QM]. Aix-Marseille Université, 2023. Français. NNT : 2023AIXM0000 . tel-04408927

HAL Id: tel-04408927

<https://hal.science/tel-04408927v1>

Submitted on 22 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

.....

THÈSE DE DOCTORAT

Soutenue à Aix-Marseille Université
le 11 décembre 2023 par

Pauline BROCHET

Étude du processus pathogénique associé à la maladie de
Chagas à l'aide d'analyses multi-omiques

Discipline

Biologie santé

Spécialité

Génomique et Bioinformatique

École doctorale

ED62

Laboratoire

INSERM U1090 – TAGC
Theories and Approaches
of Genomic Complexity

• **Composition du jury**

• Philippe NAQUET Président du jury
• Professeur universitaire - CIML

• Carl HERRMANN Rapporteur
• Professeur – IPMB

• Abdellah TEBANI Rapporteur
• MCU-PH, HDR - CHU de Rouen

• Laura CANTINI Examinatrice
• Chargée de recherche - Institut Pasteur

• Anaïs BAUDOT Invitée
• Directrice de recherche - MMG

• Christophe CHEVILLARD Directeur de thèse
• Directeur de recherche – TAGC

• Lionel SPINELLI Co-directeur de thèse
• Ingénieur de recherche – TAGC/CIML

Affidavit

Je soussignée, Pauline Brochet, déclare par la présente que le travail présenté dans ce manuscrit est mon propre travail, réalisé sous la direction scientifique de Christophe Chevillard et de Lionel Spinelli, dans le respect des principes d'honnêteté, d'intégrité et de responsabilité inhérents à la mission de recherche. Les travaux de recherche et la rédaction de ce manuscrit ont été réalisés dans le respect à la fois de la charte nationale de déontologie des métiers de la recherche et de la charte d'Aix-Marseille Université relative à la lutte contre le plagiat.

Ce travail n'a pas été précédemment soumis en France ou à l'étranger dans une version identique ou similaire à un organisme examinateur.

Fait à Marseille 16/10/2023.

A handwritten signature in black ink, appearing to read 'Brochet', with a long horizontal flourish extending to the right.

Liste de publications et participation aux conférences

Liste des publications réalisées dans le cadre du projet de thèse :

1. Laugier L, Ferreira LRP, Ferreira FM, Cabantous S, Frade AF, Nunes JP, Ribeiro RA, **Brochet P**, Teixeira PC, Santos RHB, Bocchi EA, Bacal F, Cândido DDS, Maso VE, Nakaya HI, Kalil J, Cunha-Neto E*, Chevillard C*. miRNAs may play a major role in the control of gene expression in key pathobiological processes in Chagas disease cardiomyopathy. *PLoS Negl Trop Dis.* 2020 Dec 22;14(12):e0008889. doi: 10.1371/journal.pntd.0008889. PMID: 33351798; PMCID: PMC7787679.
2. **Brochet P**, Ianni BM, Laugier L, Frade AF, Silva Nunes JP, Teixeira PC, Mady C, Ferreira LRP, Ferré Q, Santos RHB, Kuramoto A, Cabantous S, Steffen S, Stolf AN, Pomerantzeff P, Fiorelli AI, Bocchi EA, Pissetti CW, Saba B, Cândido DDS, Dias FC, Sampaio MF, Gaiotto FA, Marin-Neto JA, Fragata A, Zaniratto RCF, Siqueira S, Peixoto GL, Rigaud VO, Bacal F, Buck P, Almeida RR, Lin-Wang HT, Schmidt A, Martinelli M, Hirata MH, Donadi EA, Costa Pereira A, Rodrigues Junior V, Puthier D, Kalil J, Spinelli L*, Cunha-Neto E*, Chevillard C*. Epigenetic regulation of transcription factor binding motifs promotes Th1 response in Chagas disease cardiomyopathy. *Front Immunol.* 2022 Aug 22;13:958200. doi: 10.3389/fimmu.2022.958200. PMID: 36072583; PMCID: PMC9441916.
3. **Brochet P**, Ianni B, Nunes JPS, Frade AF, Teixeira PC, Mady C, Ferreira LRP, Kuramoto A, Pissetti CW, Saba B, Cândido DDS, Dias F, Sampaio M, Marin-Neto JA, Fragata A, Zaniratto RCF, Siqueira S, Peixoto GDL, Rigaud VOC, Buck P, Almeida RR, Lin-Wang HT, Schmidt A, Martinelli M, Hirata MH, Donadi E, Rodrigues Junior V, Pereira AC, Kalil J, Spinelli L*, Cunha-Neto E*, Chevillard C*. Blood DNA methylation marks discriminate Chagas cardiomyopathy disease clinical forms. *Front Immunol.* 2022 Sep 29;13:1020572. doi: 10.3389/fimmu.2022.1020572. PMID: 36248819; PMCID: PMC9558220.
4. **Brochet P**, Mouren JC, Hannouche L, Lopez F, Ballester B, Cunha-Neto E*, Spinelli L*, Chevillard C*. ChagasDB: 80 years of publicly available data on the molecular host response to *Trypanosoma cruzi* infection in a single database. *Database (Oxford).* 2023 May 23;2023:baad037. doi: 10.1093/database/baad037. PMID: 37221041; PMCID: PMC10205463.
5. Gallardo F*, **Brochet P***, Goudenège D, Nunes JPS, Andrieux P, Ianni B, Frade AF, Mady C, Santos RHB, Kuramoto A, Steffen S, Stolf AN, Fiorelli AI, Bocchi EA, Pissetti CW, Saba B, Dias FC, Sampaio MF, Gaiotti FA, Marin-Neto JA, Fragata A, Zaniratto RCF, Siqueira S, Peixoto GDL, Bacal F, Buch P, Almeida RR, Lin-Wang HT, Schmidt A, Hirata MH, Donadi EA, Pereira AC, Rodrigues Junior V, Martinelli M, Naslajsky M, Kalil J, Procaccio V*, Cunha-Neto E*, Chevillard C*. Mitochondrial DNA haplogroups and variants predispose to Chagas disease cardiomyopathy. **Soumis à Hearts.**
6. **Pauline Brochet**, Jorge Kalil, Vincent Procaccio, Edecio Cunha-Neto*, Lionel Spinelli*, Christophe Chevillard*. Exploring global and specific pathogenic mechanisms in Chronic Chagas Cardiomyopathy through multi-omics integration. **En préparation**, disponible sur medRxiv.

7. Nunes JPS*, Andrieux P*, **Brochet P**, Almeida RR, Kitano E, Honda AK, Iwai LK, Andrade-Silva D, Goudenège D, Alcântara Silva KD, Vieira RS, Levy D, Bydlowski SP, Gallardo F, Torres M, Bocchi EA, Mano M, Santos RHB, Bacal F, Pomerantzeff P, Laurindo FRM, Teixeira PC, Nakaya HI, Kalil J, Procaccio V, Chevillard C*, Cunha-Neto E*. Co-Exposure of Cardiomyocytes to IFN- γ and TNF- α Induces Mitochondrial Dysfunction and Nitro-Oxidative Stress: Implications for the Pathogenesis of Chronic Chagas Disease Cardiomyopathy. *Front Immunol.* 2021 Nov 11;12:755862. doi: 10.3389/fimmu.2021.755862. PMID: 34867992; PMCID: PMC8632642.
8. Silva KDA, Nunes JPS, Andrieux P, **Brochet P**, Almeida RR, Kuramoto Takara ACK, Pereira NB, Abel L, Cobat A, Zaniratto RCF, Levy D, Bydlowski SP, Ceconello I, Seguro FCBDC, Kalil J, Chevillard C*, Cunha-Neto E*. Chagas Disease Megaesophagus Patients Carrying Variant MRPS18B P260A Display Nitro-Oxidative Stress and Mitochondrial Dysfunction in Response to IFN- γ Stimulus. *Biomedicines.* 2022 Sep 7;10(9):2215. doi: 10.3390/biomedicines10092215. PMID: 36140315; PMCID: PMC9496350.

Participation aux conférences au cours de la période de thèse :

1. EMBO | EMBL Symposium : Multiomics to Mechanisms - Challenges in Data Integration.
Du 11 au 13 septembre 2019, EMBL, Heidelberg.
2. EMBO | EMBL Symposium : Systems Genetics : From Genomes to Complex Traits.
Du 29 septembre au 02 octobre 2019, EMBL, Heidelberg.
3. MarMaRa Symposium (Institut des Maladies rares).
02 juin 2022, Marseille.
Présentation orale.
4. JOBIM 2022 (Journées Ouvertes en Biologie, Informatique et Mathématiques).
Du 05 au 08 juillet 2022, Rennes.
Présentation orale.
5. COLLOQUE EDSVS 2022 (Ecole Doctorale Sciences de la vie et de la Santé - ED62).
Du 18 au 19 octobre 2022, Marseille.
Présentation orale.
6. ISMB/ECCB 2023 (31st Annual Intelligent Systems For Molecular Biology and the 22nd Annual European Conference on Computational Biology).
Du 23 au 27 juillet 2023, Lyon.

Résumé

La maladie de Chagas est une maladie parasitaire endémique d'Amérique Latine causée par *Trypanosoma cruzi*. Bien que la majorité des individus infectés demeurent asymptomatiques tout au long de leur vie, environ 30 % d'entre eux développent des formes cardiaques particulièrement graves, connues sous le nom de Cardiomyopathies Chagasiennes Chroniques (CCC), pour lesquelles aucun traitement efficace n'a encore été trouvé. Ces cardiomyopathies se caractérisent par une inflammation exacerbée du muscle cardiaque, principalement provoquée par un infiltrat immunitaire composé en grande partie de lymphocytes Th1, de lymphocytes B et de macrophages.

Afin de mieux comprendre le processus pathogénique lié aux CCC, nous avons réalisé différentes analyses omiques, incluant génomique, transcriptomique méthylomique. Ces analyses ont été effectuées sur des échantillons cardiaques prélevés chez des individus sains ainsi que chez des patients atteints de CCC. Individuellement, elles nous ont permis de caractériser les processus biologiques impliqués dans les CCC, ainsi que de confirmer l'implication d'éléments épigénétiques, incluant la méthylation de l'ADN et l'expression d'ARN non codant. Des comparaisons entre les profils épigénomiques et transcriptomiques ont confirmé leur importance dans la régulation génétique des CCC.

Ces résultats nous ont conduits à effectuer une intégration multi-omiques conduite sur ces données en utilisant l'outil MOFA+. Cette approche a mis en évidence des mécanismes communs à différents patients, tout en identifiant des aspects pathologiques spécifiques à certains d'entre eux. Cela suggère que divers processus biologiques distincts peuvent conduire au même phénotype final.

L'ensemble des études menées dans ce projet ont contribué à une meilleure compréhension du processus pathogénique associé aux CCC et ont ouvert la voie à de nouvelles hypothèses quant au développement de cette maladie.

Abstract

Chagas disease is a parasitic disease endemic to Latin America, caused by *Trypanosoma cruzi*. Although most of infected individuals stay asymptomatic during their lives, around 30% develop particularly severe cardiac forms, known as Chronic Chagasic Cardiomyopathies (CCC), for which no effective treatment has yet been found. These cardiomyopathies are characterized by an exacerbated inflammation of the heart muscle, mainly caused by an immune infiltrate mostly composed of Th1 lymphocytes, B lymphocytes and macrophages.

To gain a better understanding of the pathogenic process associated with CCC, we performed various omics analyses, including genomics, transcriptomics and methylomics. These analyses were carried out on cardiac samples taken from healthy individuals as well as CCC patients. Individually, they enabled us to characterize the biological processes involved in CCC, as well as confirming the involvement of epigenetic elements, including DNA methylation and non-coding RNA expression. Comparisons between epigenomic and transcriptomic profiles confirmed their importance in the genetic regulation associated to CCC.

These results led us to perform a multi-omics integration conducting in these omics data using the MOFA+ tool. This approach has highlighted mechanisms common to different patients, while identifying unique pathological aspects in specific cases. This suggests that various distinct biological processes can lead to the same final CCC phenotype.

Commonly, all these studies have contributed to a better understanding of CCC and opened the way to new hypotheses regarding the development of this disease.

Remerciements

Mon travail de thèse a été le fruit d'un engagement personnel important, mais n'aurait pu être mené seule. À cet égard, je tiens à exprimer ma profonde gratitude envers mes deux superviseurs de thèse, Christophe Chevillard et Lionel Spinelli. Vous avez toujours été présents pour moi, que ce soit dans l'aide apportée à mon travail ou par vos nombreux encouragements dans l'exploration de nouvelles approches. Votre présence a été précieuse non seulement sur le plan professionnel, mais aussi personnel, à travers de multiples discussions et conseils m'ayant apporté un grand soutien moral.

Christophe, c'est toi qui m'as incitée à entreprendre cette thèse, une idée qui ne me paraissait initialement pas envisageable. Dès mon stage de M1, tu as su progressivement me convaincre de la valeur et de l'intérêt de ce projet. Depuis 2017, toutes tes promesses ont été tenues (et malgré le Covid !), que ce soit dans la participation de congrès comme dans la mise en place de collaborations et d'échanges internationaux. Des vents froids d'Allemagne aux plages chaudes du Brésil, ces moments de partage resteront gravés dans ma mémoire.

Lionel, lorsque tu as rejoint mon projet de thèse, ta réputation d'excellence te précédant m'a quelque peu intimidée. 4 ans plus tard, je suis toujours aussi intimidée, mais très heureuse que tu aies accepté. Tu m'as fait comprendre que derrière toute méthode mathématique, des espaces vectoriels se cachent... Mais assez plaisanté, ton expertise a été indispensable et a considérablement enrichi mes compétences et ma rigueur professionnelle. Au-delà du cadre professionnel, nos nombreuses discussions m'ont fait évoluer personnellement et m'aideront certainement dans le futur à prendre de sages décisions.

Pendant mon doctorat, vous avez tous deux été complémentaires (bon flic mauvais flic, je vous laisse deviner qui est qui), apportant un soutien précieux à divers aspects de mon travail, et j'espère que nos échanges perdureront à l'avenir. Je n'aurais pas pu rêver d'un meilleur encadrement, et c'est donc avec une certaine tristesse que je clôture cette thèse, notamment en mettant fin à notre collaboration.

Je remercie mon jury, Philippe Naquet, Carl Herrmann, Abdellah Tebani, Anaïs Baudot et Laura Cantini, pour avoir accepté d'évaluer mon travail. Je remercie également les membres de mon comité de thèse, Anaïs Baudot, Julien Barc, Benoît Ballester et Fayrouz Hammal, pour avoir veillé à la bonne progression de ma thèse.

Je souhaite également remercier Pauline Andrieux. Ton arrivée dans l'équipe a été une agréable surprise (après m'être assurée que tu appréciais le Seigneur des Anneaux, bien sûr). Initialement, j'appréciais mon confort solitaire au bureau, mais tu m'as rapidement fait changer d'avis et ta présence m'est devenue essentielle. Au cours de ces 4 ans, tu m'as soutenue et inspirée par ta force et ton courage face aux différentes épreuves de la vie. Je te souhaite tout le meilleur pour l'avenir, mais je sais que tu réussiras quelle que soit la difficulté. Que la mitochondrie soit avec toi !

Je tiens à remercier toutes les personnes qui ont contribué à mon projet. Je remercie Delphine Potier et Pierre Milpied pour leur aide et leurs conseils sur différentes analyses. Je suis reconnaissante envers le groupe MaBios pour m'avoir permis de présenter mes travaux lors de leurs séminaires externes et pour leurs retours constructifs. Je tiens à remercier Benoît Ballester et Jean-Christophe Mouren pour leur contribution au projet ChagasDB, ainsi que Quentin Ferré pour ses conseils pendant mes analyses.

Merci à l'ensemble des membres du TAGC, et notamment le directeur de l'unité, Pascal Rihet, pour m'avoir accueillie au sein du laboratoire. Je remercie tous mes collègues pour ces bons moments passés, avec une pensée particulière pour Florian et Héliéa. Je remercie également les enseignants-chercheurs du TAGC, spécialement Andreas Zanzoni, qui ont été d'une grande aide pour mes enseignements tout au long de mes quatre années de thèse.

Je souhaite exprimer ma reconnaissance envers nos collaborateurs brésiliens pour leurs conseils, ainsi que pour leur chaleureux accueil au Brésil. Merci à Edecio Cunha-Neto pour sa précieuse expertise en tant que spécialiste de la maladie de Chagas. João Paulo Nunes, merci d'avoir été un guide exceptionnel pendant mon séjour au Brésil, et Amanda Farage Frade, merci pour ton hospitalité, je penserai à toi quand j'irai à la fruiterie ! Nous avons partagé des moments merveilleux.

Je remercie Cameron MacKereth, Gisèle Bronner, Hélène Gardon et Bianca Habermann pour m'avoir accueillie dans leurs différentes équipes, m'offrant ainsi l'opportunité de découvrir le monde de la recherche et de me guider vers ma situation actuelle. Je tiens particulièrement à remercier Arantxa Etchechurry pour m'avoir initiée à la génétique et transmis sa passion des sciences de la vie.

Je tiens à remercier chaleureusement mes proches pour leur soutien indéfectible. Je remercie ma famille pour m'avoir permis de mener mes études, mais aussi d'avoir été un pilier constant dans les moments difficiles. Vous vous êtes toujours intéressés à mon travail, malgré vos théories parfois farfelues (les mitochondries !!!). Je remercie mes frères d'avoir été là pour moi, et mes parents pour m'avoir toujours fait confiance et soutenue dans mes choix. Je remercie également mes amis, Ange, Maialen, Soizic et Quentin, pour votre soutien continu depuis le début de mes études. Votre amitié m'a toujours été très précieuse.

Enfin, je remercie évidemment ~~mes chats~~ mon compagnon, Laurent Hannouche, qui a été à mes côtés dès mon arrivée au TAGC et qui m'a apporté un soutien inestimable tout au long de ma thèse. Tu t'es constamment soucié de mon bien-être, veillant à ce que je me nourrisse correctement et que je prenne du repos. Ton engagement s'est également reflété dans mon travail, en consacrant tes vacances à m'aider à créer ChagasDB, un projet qui me tenait énormément à cœur. Tu as également été présent pour relire mes articles, ma thèse et pour écouter patiemment mes présentations orales. Merci pour tout.

C'est grâce à la contribution de chacun d'entre vous que j'ai pu mener à bien mes études et mon projet de thèse, et que j'envisage avec sérénité vers mon avenir professionnel et personnel. Je vous adresse mes plus sincères remerciements.

Liste des abréviations

ACP : Analyse en Composantes Principales

ADN : Acide Désoxyribonucléique

APC : Cellules présentatrices d'antigènes

ARN : Acide Ribonucléique

ARNcirc : ARN circulaire

ARNInc : ARN long non codant

ARNr : ARN ribosomique

ARNt : ARN de transfert

ASY : Asymptomatique

ATAC-Seq : Assay for Transposase-Accessible Chromatin Using Sequencing

ATP : Acide Adénosine-Triphosphorique

AVC : Accident Vasculaire Cérébral

BCR : B Cell Receptor

BDD : Base De Données

Ca²⁺ : Calcium

CCC : Cardiomyopathie(s) Chagasiq(u)e(s) Chroniq(u)e(s)

CDC : Centers for Diseases Control

Cellules NK : Cellules Natural-Killer

ChIP-Seq : Chromatin Immunoprecipitation followed by Sequencing

CMD : Cardiomyopathie Dilatée

DALY : Disability-Adjusted Life Years

DE : Différentiellement Exprimé

DEG : Differentially Expressed Genes

DIABLO : Data Integration Analysis for Biomarker discovery using Latent cOmponents

DTU : Discrete Typing Units

ECG : Electrocardiogramme

ELISA : Enzyme-Linked Immunosorbent Assay

FAIR : Findable, Accessible, Interoperable, Reusable

FT : Facteur de Transcription

GIPL : Glycoinositolphospholipides libre
GO : Gene Ontology
GPI : Glycosylphosphatidylinositol
GWAS : Genome-Wide Association Study
HA : Hemagglutination Assay
HLA : Human Leucocyte Antigen
HMDD : Human MicroRNA Disease Database
IFAT : Indirect Fluorescent Antibody Test
IFN- γ : Interferon gamma
Ig : Immunoglobuline
IL-x : Interleukine x
IP3 : 1,4,5-inositol triphosphate
JIVE : Joint and Individual Variation Explained
KEGG : Kyoto Encyclopedia of Gene and Genomes)
KNNOR : K-Nearest Neighbor OverSampling approach
LB : Lymphocyte B
LT : Lymphocyte T
LVEF : Left Ventricular Ejection Fraction
Lymphocyte Th : Lymphocyte T helper
Lymphocyte Treg : Lymphocyte T régulateur
MFA : Multiple Factor Analysis
MHC : Major Histocompatibility Complex
miARN : Micro-ARN
MICA : Multiple Co-Inertia Analysis
MI-MFA : Multiple Imputation in Multiple Factor Analysis
miR-x : Micro-ARN x
ML : Machine Learning
MOFA : Joint and Individual Variation Explained
mRMR : Maximum Relevance and Minimum Redudancy
MyD88 : Myeloid differentiation primary response 88
NLR : NOD-Like Receptor

NYHA : New York Heart Association
OMIM : Online Mendelian Inheritance in Man
OMS : Organisation Mondiale de la Santé
PAHO : Pan American Health Organization
PAMP : Pathogen Associated Molecular Patterns
PARADIGM : PATHway Recognition Algorithm using Data Integration on Genomic Models
PCR : Polyclonal Chain Reaction
PID : Pathway Interaction Database
PRR : Pattern Recognition Receptor
RFE : Recursive Feature Elimination
RISC : RNA-Induced Silencing Complex
RNS : Reactive Nitrogen Species
ROS : Reactive Oxygene Species
SMOTE : Synthetic Minority Over-sampling Technique
SNP : Single Nucleotide Polymorphism
T. cruzi : *Trypanosoma cruzi*
TCR : T Cell Receptor
TLR : Toll-Like Receptor
TNF- α : Tumor necrosis factor alpha
TOBMI : Trans-Omic Block Missing Data Imputation
TSS : Transcription Start Site
WES : Whole-Exome Sequencing
WGS : Whole-Genome Sequencing
WHO : Wold Health Organization

Table des figures

| | |
|---|----|
| Figure 1. Carlos Chagas avec Berenice. | 20 |
| Figure 2. Premières illustrations de <i>T. cruzi</i> réalisées par Carlos Chagas en 1909..... | 20 |
| Figure 3. Les différentes formes de <i>T. cruzi</i> (barres d'échelle : 5µm)..... | 22 |
| Figure 4. Le cycle de vie de <i>Trypanosoma cruzi</i> | 24 |
| Figure 5. Distribution des différentes souches de <i>T. cruzi</i> en Amérique Latine, en fonction de leur reproduction domestique ou sylvatique..... | 26 |
| Figure 6. Estimation du nombre de personnes atteintes de la maladie de Chagas à travers le monde en 2018. | 28 |
| Figure 7. Distribution de la valeur DALY (pour 100.000 habitants) à travers le Brésil en 2016. | 29 |
| Figure 8. Estimation du nombre d'individus atteints de la maladie de Chagas en Europe. | 30 |
| Figure 9. Photographies montrant un chagome (A) et le signe de Romana (B). | 32 |
| Figure 10. Échantillon histologique du ventricule droit d'un chien infecté par <i>T. cruzi</i> | 34 |
| Figure 11. Courbe de survie temporelle des patients atteints de CCC et d'autres pathologies cardiaques. | 35 |
| Figure 12. Photographie de la forme trypomastigote de <i>T. cruzi</i> présent dans un échantillon sanguin coloré au giemsa. | 37 |
| Figure 13. Représentation schématique du complexe QRS observé sur un ECG normal. | 39 |
| Figure 14. Comparaison d'un ECG normal (A) avec un ECG de patient atteint de CCC (B). La diminution du voltage QRS et l'amplitude de l'onde T sont visibles. | 39 |
| Figure 15. Pénétration de <i>T. cruzi</i> dans les cellules par phagocytose..... | 46 |
| Figure 16. Pénétration active de <i>T. cruzi</i> dans des cellules non phagocytaires A) par le biais des lysosomes ou B) en utilisant la membrane cellulaire pour former une vacuole. | 47 |
| Figure 17. Activation de la réponse immunitaire innée via les TLR dans un macrophage. | 48 |
| Figure 18. Activation de la réponse immunitaire innée indépendamment des TLR. | 49 |
| Figure 19. Mise en place des méthodes de destruction du parasite par les cellules phagocytaires. | 50 |
| Figure 20. Représentation d'un ganglion lymphatique. | 52 |
| Figure 21. Représentation simplifiée de l'activation des lymphocytes B et T dans les ganglions lymphatiques. | 54 |
| Figure 22. Représentation simplifiée de la réponse immunitaire adaptative contre <i>T. cruzi</i> | 57 |
| Figure 23. Déséquilibre dans la réponse pro- et anti-inflammatoire chez les patients atteints de CCC par rapport aux patients ASY. | 60 |
| Figure 24. Schéma récapitulant les connaissances actuelles concernant le rôle de la réponse immunitaire dans le développement des lésions cardiaques observées dans les CCC..... | 62 |
| Figure 25. Présentation des principales "omiques" permettant d'analyser la globalité d'un processus biologique. | 65 |
| Figure 26. Impact de la méthylation de l'ADN (violet) dans la régulation de l'expression génique. | 69 |
| Figure 27. Schéma représentant le mode de fonctionnement des miARN dans la régulation d'un gène cible. | 70 |
| Figure 28. Évolution du nombre de nouvelles bases de données publiées dans la revue NAR (Nucleic Acide Research) entre 1991 et 2016. | 71 |
| Figure 29. Comparaison des profils d'expression entre les contrôles sains, les CCC et les CMD. | 80 |
| Figure 30. Réseau d'interaction entre les miARN dérégulés chez les CCC et leurs cibles potentielles. | 82 |

| | |
|---|----|
| Figure 31. Réseau d'interactions prédites in silico entres les miARN et les gènes dérégulés chez la souris 30 jours post-infection par T. cruzi. | 84 |
| Figure 32. Représentation artistique de la complexité des analyses multi-omiques. | 86 |
| Figure 33. Evolution du nombre de publications recensées sur Pubmed incluant le terme "multiomic". | 88 |
| Figure 34. Représentation schématique de la différence entre l'intégration tardive (late fusion) et l'intégration précoce (early fusion) de données provenant de différentes sources. | 89 |
| Figure 35. Illustration des trois types de méthodes détaillées ci-dessous. | 90 |
| Figure 36. Représentation schématique du principe de la factorisation matricielle. | 92 |
| Figure 37. Représentation schématique du fonctionnement des méthodes SMOTE et KNNOR. | 97 |

Table des matières

| | |
|---|-----------|
| AFFIDAVIT | 2 |
| LISTE DE PUBLICATIONS ET PARTICIPATION AUX CONFERENCES | 3 |
| LISTE DES PUBLICATIONS REALISEES DANS LE CADRE DU PROJET DE THESE : | 3 |
| PARTICIPATION AUX CONFERENCES AU COURS DE LA PERIODE DE THESE : | 5 |
| RESUME | 6 |
| ABSTRACT | 7 |
| REMERCIEMENTS | 8 |
| LISTE DES ABREVIATIONS | 11 |
| TABLE DES FIGURES | 14 |
| TABLE DES MATIÈRES | 16 |
| INTRODUCTION | 19 |
| CHAPITRE 1 : INTRODUCTION À LA MALADIE DE CHAGAS | 19 |
| 1.1. <i>Découverte de la maladie</i> | 19 |
| 1.2. <i>La transmission de la maladie</i> | 22 |
| 1.2.1. Les différents hôtes | 22 |
| 1.2.2. Le cycle de vie du parasite | 22 |
| 1.2.3. Les autres modes de transmission..... | 24 |
| 1.3. <i>De la diversité du parasite à la diversité des symptômes</i> | 26 |
| 1.4. <i>Épidémiologie</i> | 27 |
| CHAPITRE 2 : L'ÉVOLUTION DE LA MALADIE | 32 |
| 2.1. <i>La phase aiguë</i> | 32 |
| 2.2. <i>La phase chronique</i> | 33 |
| 2.2.1. Les complications cardiaques : la Cardiomyopathie Chagastique Chronique | 33 |
| 2.2.2. Les complications gastro-intestinales | 35 |
| 2.3. <i>Diagnostic</i> | 36 |
| 2.3.1. Evaluation de la contamination par <i>T. cruzi</i> | 36 |
| 2.3.2. Diagnostic des cardiomyopathies chagastiques chroniques..... | 38 |
| 2.3.2. Diagnostic des atteintes digestives chroniques | 40 |
| 2.3.3. L'utilisation de biomarqueurs spécifiques des différentes formes de la maladie | 41 |
| 2.4. <i>Traitements</i> | 42 |
| 2.4.1. Les traitements antiparasitaires..... | 42 |
| 2.4.2. Les traitements des atteintes cardiaques..... | 43 |
| 2.4.3. Les traitements des atteintes digestives..... | 44 |
| CHAPITRE 3. DE LA DÉFENSE DE L'HÔTE AU DÉVELOPPEMENT DES CARDIOMYOPATHIES CHAGASTIQUES CHRONIQUES | 45 |
| 3.1. <i>L'infection de l'hôte par le parasite et le déclenchement de la réponse immunitaire innée</i> | 45 |

| | |
|--|------------|
| 3.1.1. L'entrée du parasite dans les cellules | 45 |
| 3.1.2. Le déclenchement de la réponse immunitaire innée | 47 |
| 3.1.3. La destruction du parasite | 49 |
| 3.2. <i>La réponse immunitaire adaptative</i> | 51 |
| 3.2.1. Le déclenchement de la réponse immunitaire adaptative | 51 |
| 3.2.2. La réponse immunitaire adaptative contre <i>T. cruzi</i> | 54 |
| 3.3. <i>L'évasion de T. cruzi face à la réaction immunitaire</i> | 57 |
| 3.4. <i>Le développement des formes cardiaques</i> | 59 |
| 3.4.1. Une absence de régulation de la réaction immunitaire..... | 59 |
| 3.4.2. Une réaction auto-immune..... | 62 |
| CHAPITRE 4 : ETUDE DE LA MALADIE DE CHAGAS PAR L'UTILISATION DE DIFFÉRENTES OMIQUES | 65 |
| 4.1. <i>Qu'est-ce que les "omiques" ?</i> | 65 |
| 4.1.1. Les molécules jouant un rôle actif dans les processus biologiques | 66 |
| 4.1.2. Les éléments génétiques et épigénétiques impliqués dans les processus de régulation | 67 |
| 4.2. <i>Les bases de données pour l'interprétation des données omiques</i> | 71 |
| 4.2.1. Les bases de données de fonctions | 73 |
| 4.2.2. Les bases de données d'interactions..... | 74 |
| 4.3. <i>L'utilisation des omiques dans la compréhension des Cardiomyopathies Chagasiques Chroniques</i> | 75 |
| 4.3.1. Découverte d'une contribution génétique aux cardiomyopathies chagasiques chroniques | 75 |
| 4.3.2. Identification et caractérisation des processus pathogéniques associés aux cardiomyopathies chagasiques chroniques..... | 78 |
| 4.3.3. Recherche des causes de dérégulations de Chagas..... | 81 |
| CHAPITRE 5 : LES ANALYSES MULTI-OMIQUES | 86 |
| 5.1. <i>Introduction aux analyses multi-omiques</i> | 86 |
| 5.1.1. Les débuts du multi-omiques | 87 |
| 5.1.2. L'intégration précoce..... | 89 |
| 5.2. <i>Les difficultés techniques liées à l'intégration multi-omiques</i> | 94 |
| 5.2.1. Les échantillons manquants..... | 94 |
| 5.2.2. La malédiction de la dimension | 95 |
| 5.2.3. L'hétérogénéité des données | 98 |
| 5.2.4. Le choix de la méthode..... | 98 |
| 5.3. <i>La complexité associée à l'interprétation des intégrations multi-omiques</i> | 100 |
| CHAPITRE 6 : MON PROJET DE THÈSE | 103 |
| RESULTATS | 105 |
| 1. LE PROFIL GÉNÉTIQUE DES MITOCHONDRIES EST ASSOCIÉ AU DÉVELOPPEMENT DES CARDIOMYOPATHIES CHAGASIQUES CHRONIQUES..... | 105 |
| 2. DIFFÉRENTES VARIABLES ÉPIGÉNÉTIQUES ONT UN IMPACT DANS LE PROCESSUS PATHOGÉNIQUE ASSOCIÉ AUX CARDIOMYOPATHIES CHAGASIQUES CHRONIQUES | 129 |
| 2.1. <i>Des miARN sont associés aux cardiomyopathies chagasiques chroniques, et semblent réguler des fonctions biologiques essentielles au développement de cette pathologie</i> | 129 |

| | |
|--|------------|
| 2.2. <i>La méthylation de l'ADN est impliquée dans le développement des cardiomyopathies chagasiques chroniques, et semble réguler l'expression de gènes en affectant la fixation de facteurs de transcription</i> | 154 |
| 2.2.1. <i>Implication de la méthylation sur la régulation génique et les processus biologiques dans les cardiomyopathies chagasiques chroniques</i> | 154 |
| 2.2.2. <i>Confirmation in silico de l'impact de la méthylation de l'ADN dans le développement des cardiomyopathies chagasiques chroniques</i> | 180 |
| 3. EXPLORATION MULTI-OMIQUES DES CARDIOMYOPATHIES CHAGASIQUES CHRONIQUES : DE LA COLLECTE DE DONNÉES À L'INTÉGRATION APPROFONDIE | 189 |
| 3.1. <i>Centralisation des résultats omiques de la maladie de Chagas dans une seule base de données</i> | 189 |
| 3.2. <i>L'intégration multi-omiques a mis en évidence des mécanismes pathologiques uniques chez des patients atteints de cardiomyopathie chagastique chronique</i> | 197 |
| DISCUSSION | 234 |
| 1. DES DIFFÉRENCES GÉNÉTIQUES ET ÉPIGÉNÉTIQUES ASSOCIÉES AUX CARDIOMYOPATHIES CHAGASIQUES CHRONIQUES | 235 |
| 1.1. <i>Des gènes aux processus biologiques</i> | 235 |
| 1.1.1. <i>Les gènes codants</i> | 235 |
| 1.1.2. <i>Les ARN non codants</i> | 237 |
| 1.2. <i>L'implication de la méthylation dans le développement des cardiomyopathies chagasiques chroniques</i> | 238 |
| 1.3. <i>La génétique mitochondriale dans le développement des cardiomyopathies chagasiques chroniques</i> | 239 |
| 1.4. <i>Une vue d'ensemble des analyses omiques dans la compréhension des cardiomyopathies chagasiques chroniques</i> | 241 |
| 1.5. <i>Limites et perspectives des analyses simple omiques</i> | 242 |
| 2. COMPRENDRE LES RÉGULATIONS GÉNÉTIQUES GRÂCE À L'INTÉGRATION MULTI-OMIQUES | 244 |
| 2.1. <i>L'intégration tardive</i> | 245 |
| 2.2. <i>L'intégration tardive</i> | 247 |
| 2.2.1. <i>La complexité du pré-traitement des données</i> | 247 |
| 2.2.2. <i>Le choix de la méthode</i> | 250 |
| 2.2.3. <i>L'interprétation des résultats</i> | 250 |
| 2.2.4. <i>Limites et perspectives</i> | 253 |
| 3. CONTRIBUTION DE CETTE THÈSE DANS LA RECHERCHE POUR LA MALADIE DE CHAGAS | 255 |
| 3.1. <i>Des progressions dans la compréhension du processus pathogénique associé aux cardiomyopathies chagasiques chroniques</i> | 255 |
| 3.2. <i>Au-delà de la maladie de Chagas : des approches bio-informatiques avancées</i> | 257 |
| 3.3. <i>De l'analyse in silico à une application médicale</i> | 258 |
| REFERENCES | 259 |
| ANNEXES | 280 |
| ANNEXE 1 | 281 |
| ANNEXE 2 | 307 |

INTRODUCTION

Chapitre 1 : Introduction à la maladie de Chagas

La maladie de Chagas est une maladie parasitaire endémique d'Amérique Latine causée par *Trypanosoma cruzi*, un protozoaire. Ce chapitre a pour but de présenter le parasite, les modes de contamination et les populations touchées par cette maladie.

1.1. Découverte de la maladie

La maladie de Chagas a été découverte en 1909 par Carlos Ribeiro Justiniano Chagas. Né en 1878 dans l'État de Minas Gerais au Brésil, Carlos Chagas a obtenu son diplôme de médecine en 1902 (1). Il a alors commencé sa thèse sur les aspects hématologiques du paludisme et en 1905, il mène une première campagne de prophylaxie au Brésil. En 1906, il rejoint l'institut de Oswaldo Cruz, qui le charge un an plus tard de mener une seconde campagne prophylactique anti paludisme dans l'Etat de Minas Gerais. Lors de cette mission, il apprend l'existence de la punaise *Triatoma infestans*. Aussi appelée "barbeiro", "kissing bug" ou "barber bug", elle pique majoritairement le visage pour prendre un repas sanguin. Il découvre également une nouvelle espèce de Trypanosome, qu'il nomme *cruzi*, en l'honneur de son mentor Oswaldo Cruz. D'abord identifié dans le sang d'un chat, il le retrouve peu de temps après chez un humain. En effet, le 14 février 1909, il consulte une patiente de deux ans, Berenice, atteinte de fièvre élevée, une hépatosplénomégalie (augmentation du volume du foie), un œdème facial et des parasites dans le sang (**figure 1**). Il cultive alors le parasite chez des mammifères et conclut la même année ses premiers rapports sur le cycle de vie du parasite, son infection chez l'homme et les symptômes en découlant (**figure 2**).



Figure 1. Carlos Chagas avec Berenice.

Adaptée de Chen Chao M.D. et al, 2020 (2).

Deux ans plus tard, en 1911, il publie la description finale du parasite, pathologie et épidémiologie (3). Il est nommé une première fois pour recevoir le prix Nobel de médecine en 1913, qui sera finalement attribué à Charles Robert Richet.

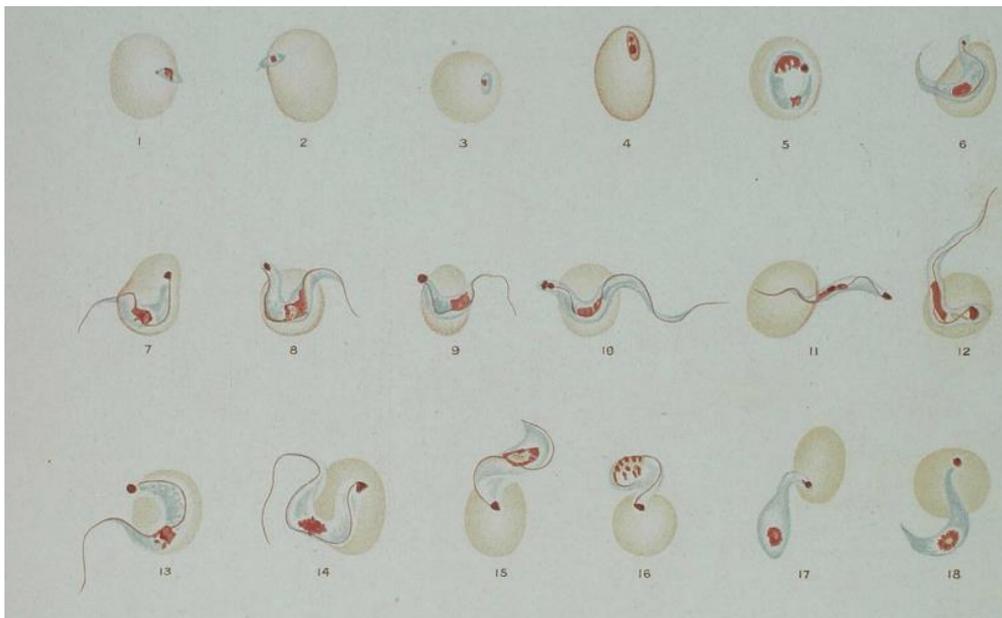


Figure 2. Premières illustrations de T. cruzi réalisées par Carlos Chagas en 1909.

Adaptée de Nova tripanozomiaze humana, 1909.

Les années qui suivent, Carlos Chagas se penche en détail sur les atteintes cardiaques dues à l'infection par *T. cruzi*. Il commence par remarquer des épisodes d'arythmie particulièrement fréquents chez les individus ayant moins de 50 ans dans les régions endémiques, ainsi que des nids de parasites entourés d'importants infiltrats monocitaires à l'autopsie. Il note également l'apparition d'insuffisance cardiaque et de fibrose (4). En 1922, Carlos Chagas a totalement caractérisé les Cardiomyopathies Chagasiques Chroniques (CCC) (5).

À partir de 1916, en réponse à la renommée nationale et internationale que rencontre Carlos Chagas, des détracteurs apparaissent et critiquent ses travaux. La difficulté à identifier des individus infectés par *T. cruzi* à travers l'Amérique Latine, et la différence des symptômes provoqués par la Trypanosomiase Américaine contrairement aux Trypanosomiasés Européennes apportent des doutes sur ses découvertes. Plus impactant encore, d'autres chercheurs n'ayant pas atteint sa position (nommé Directeur de la Recherche Publique en 1920) dénigrent ses travaux, entachant sa réputation et sa crédibilité. Il est à nouveau nominé pour le prix Nobel de médecine en 1921, mais ne l'obtiendra pas. En effet, aucun candidat n'a reçu le prix cette année-là, le comité estimant qu'aucune découverte nominée ne répondait aux critères d'exigences du prix Nobel (6). Toutes ces détractations font tomber la maladie dans l'oubli, qui n'est plus enseignée à l'école de médecine, et donc rarement diagnostiquée. C'est grâce à une amélioration apportée au test de détection de *T. cruzi* en 1936, le rendant plus stable, que la caractérisation de la maladie, notamment chez des jeunes ayant des arythmies sévères, a repris (4).

En 1961, âgée de 53 ans, Berenice a été retrouvée vivante dans l'État de Minas Gerais, totalement asymptomatique. En revanche, le parasite a bien été détecté dans son sang, et les trypanosomes cultivés à partir de son sang sont toujours utilisés en tant que la souche "Berenice". Berenice est décédée à 73 ans de causes autres que la maladie de Chagas, après être restée asymptomatique toute sa vie.

1.2. La transmission de la maladie

1.2.1. Les différents hôtes

Le parasite responsable de la maladie de Chagas, *Trypanosoma cruzi*, appartient à l'ordre des protozoaires, tout comme *Trypanosoma brucei* (maladie du sommeil), *Plasmodium falciparum* (paludisme), *Toxoplasma gondii* (toxoplasmose) ou *Leishmania* (leishmaniose) (7). Il s'agit d'un organisme eucaryote unicellulaire qui évolue à travers 2 hôtes : un triatome et un mammifère.

Le triatome est une punaise appartenant à l'ordre des reduviidae. La majorité des 130 espèces de triatome sont hématophages, et peuvent potentiellement transmettre *T. cruzi*. Cependant, du fait de sa proximité avec les humains (il vit dans les habitations en zone rurale), *Triatoma infestans* est le principal vecteur de la maladie de Chagas (8).

Concernant les mammifères, de nombreuses espèces peuvent être infectées et servir de réservoir (9). Tous les ordres sont concernés : les chiroptères, les primates, les marsupiaux, les carnivores... Cela s'applique évidemment aux animaux domestiques, comme les chèvres, les chiens ou les chats.

1.2.2. Le cycle de vie du parasite

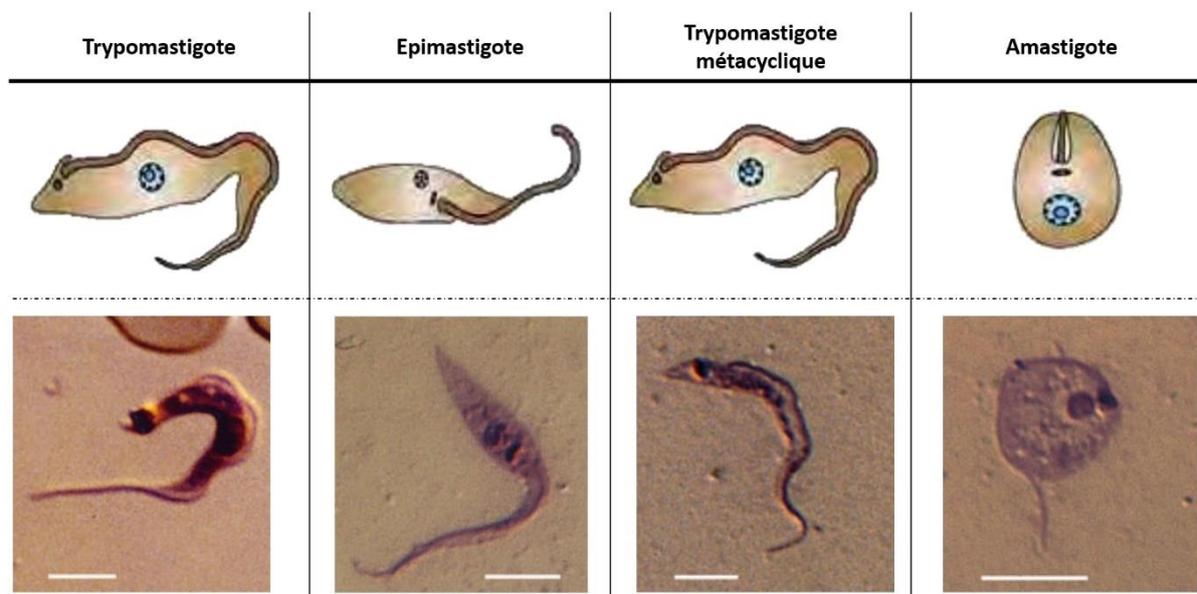


Figure 3. Les différentes formes de *T. cruzi* (barres d'échelle : 5µm).

Adapté de Fontes A. et al., 2011 (10) et Meiser CK. et al., 2011 (11).

Le cycle de vie de *T. cruzi* peut être divisé en deux phases : chez l'insecte vecteur, et chez le mammifère. Lorsque le triatome hématophage pique un mammifère pour son repas sanguin, il va ingérer le parasite présent dans le sang sous sa forme trypomastigote (allongée, avec un long flagelle lui permettant de se déplacer) (**figure 3**). Ils deviennent des épimastigotes dans son intestin, et prolifèrent par scissiparité (reproduction asexuée par simple division de l'organisme). Ils migrent alors dans l'intestin postérieur, et se différencient en trypomastigotes métacycliques.

Lorsque le triatome pique à nouveau un mammifère, les trypanosomes métacycliques présents dans ses intestins peuvent être excrétés dans ses fèces. Ces derniers peuvent ensuite entrer dans le corps du mammifère à travers les muqueuses, ou à travers la plaie lorsque le mammifère gratte la zone piquée. Une fois à l'intérieur de l'organisme, les parasites interagissent avec les récepteurs présents à la surface des cellules de l'hôte, ce qui déclenche le processus d'internalisation. Les trypanosomes créent un compartiment intracellulaire spécifique, connu sous le nom de vacuole parasitophore, dans laquelle ils subissent une transformation en une forme arrondie appelée amastigote. Cette transformation se produit pour échapper au système immunitaire et pour se développer en toute sécurité à l'intérieur de la cellule hôte. Les parasites se répliquent alors par scissiparité jusqu'à ce que les cellules soient remplies de réplicats. Ils s'allongent, développent à nouveau un long flagelle, se différencient en trypomastigotes non réplicatifs et lysent la membrane cellulaire. Ils envahissent alors les cellules adjacentes pour recommencer un cycle de réplication, ou circulent dans le sang et la lymphe et se disséminent dans l'organisme (12) (**figure 4**). De récentes études ont cependant démontré que les formes amastigotes (dans les cellules des mammifères) et les formes épimastigotes (dans l'intestin du rédve) semblent parfois effectuer des échanges génétiques, remettant en question le modèle de reproduction purement clonal du parasite, et introduisant la possibilité d'une reproduction sexuée (13).

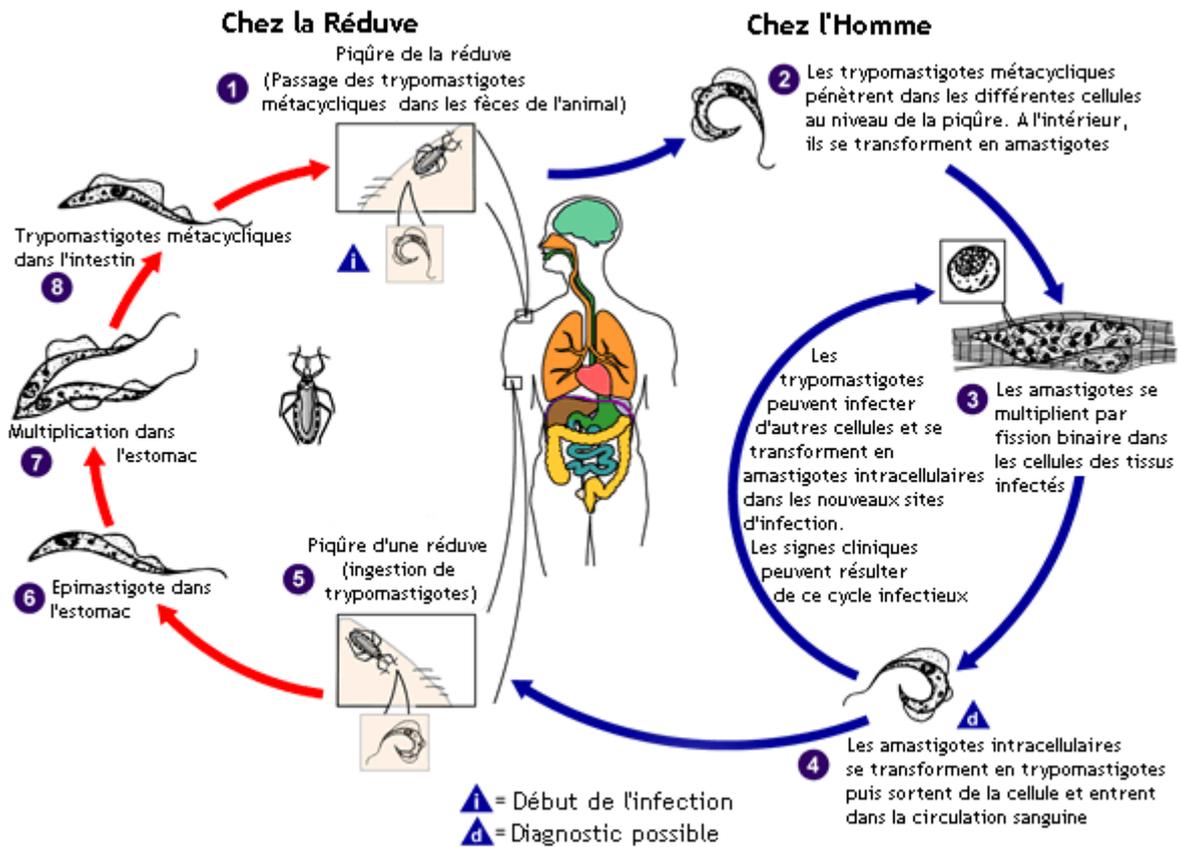


Figure 4. Le cycle de vie de *Trypanosoma cruzi*.

Adapté du CDC (Centers for Diseases Control and Prevention).

1.2.3. Les autres modes de transmission

La contamination vectorielle, telle que présentée dans le cycle de vie classique du parasite, représente la grande majorité des cas de contamination en zone endémique. Toutefois, 4 autres modes de transmissions alternatifs existent : sanguine, congénitale, greffe d'organe et orale.

- La transmission sanguine

En 1960, selon les estimations de l'OMS, environ 7 millions de nouvelles contaminations ont eu lieu par transfusion sanguine à travers l'ensemble de l'Amérique latine, dont 6.000 à 10.000 cas à Rio de Janeiro et São Paulo. Entre 1961 et 1963, une étude a montré que, sur 4.595 donneurs de sang, 1.8% étaient contaminés. Aujourd'hui, au Brésil, comme dans de nombreux pays (y compris la France), les poches de sang sont testées, mais ce n'est pas systématique dans tous les pays (ex. Bolivie, Italie).

Ce problème peut parfois être très important, comme dans les villes Santa Cruz de la Sierra et Cochabamba (Bolivie), où 1/4 des poches de sang peuvent être contaminées (14).

- La transmission congénitale

T. cruzi, circulant dans le sang de l'organisme, peut être transmis de manière congénitale. Environ 5% des femmes infectées vont propager le parasite à leur fœtus. Ce dernier est le plus souvent asymptomatique, mais peut présenter les symptômes d'une infection aiguë, comme toute personne récemment contaminée. Environ 9.000 nouveau-nés sont contaminés chaque année en Amérique Latine. En dehors de la zone endémique, la majorité des nouvelles contaminations sont dues à une transmission congénitale (15). Ce mode de transmission peut être répété à chaque grossesse, et transmis à travers les générations, devenant de ce fait de plus en plus important (16).

- La transmission par transplantation d'organe

Des cas de transmission ont été observés à travers des organes provenant de donneurs chroniquement infectés en Argentine, au Brésil ou aux États-Unis, avec des taux d'infection variables en fonction de l'organe. Alors qu'ils restent modérés dans le contexte des transplantations de foie ou de reins (13 à 22 % d'infection après transplantation), le risque de transmission est nettement plus élevé dans le cas des transplantations cardiaques (pouvant atteindre jusqu'à 75 %)(17).

- La transmission orale

La transmission orale se fait par ingestion de nourriture contaminée. Il peut s'agir d'aliments contaminés par des excréments de triatomés, aspergés par des sécrétions de glandes anales odoriférantes, préparés avec des triatomés écrasés (comme le jus de canne à sucre) ou encore de viandes d'animaux infectés insuffisamment cuites. Le parasite entre alors dans la muqueuse gastrique, où il prend sa forme amastigote, se réplique et poursuit son cycle comme décrit précédemment (18). Ce mode de contamination induit souvent une épidémie locale. Par exemple, en 2005, 25 individus ont été contaminés à Santa Catarina (Brésil) après avoir ingéré du jus de canne à sucre, et 3 en sont décédés (19).

1.3. De la diversité du parasite à la diversité des symptômes

Trypanosoma cruzi constitue une espèce à part entière, qui abrite une variété de souches distinctes. Par convention, ces souches sont désignées par l'acronyme DTU (Discrete Typing Units). 6 souches différentes ont été identifiées, allant de TcI à TcVI (20). Une 7ème souche, appelée Tcbat, a également été découverte (21). TcI, TcII, TcV et TcVI sont les souches les plus souvent responsables des infections humaines, infectant majoritairement les espèces domestiques (animales et humaines). A l'inverse, TcIII et TcIV contaminent en grande partie les espèces sauvages (sylvatiques). Leurs répartitions géographiques se démarquent également. TcI, par exemple, occupe la majeure partie du territoire sud-américain, notamment en Colombie et au Venezuela. TcIV est également présente dans le nord et le centre de l'Amérique du Sud, tandis que TcV et TcVI se concentrent davantage vers le sud. TcII et TcIII, quant à elles, sont plus fortement représentées dans le centre de l'Amérique latine, en particulier au Brésil (figure 5).

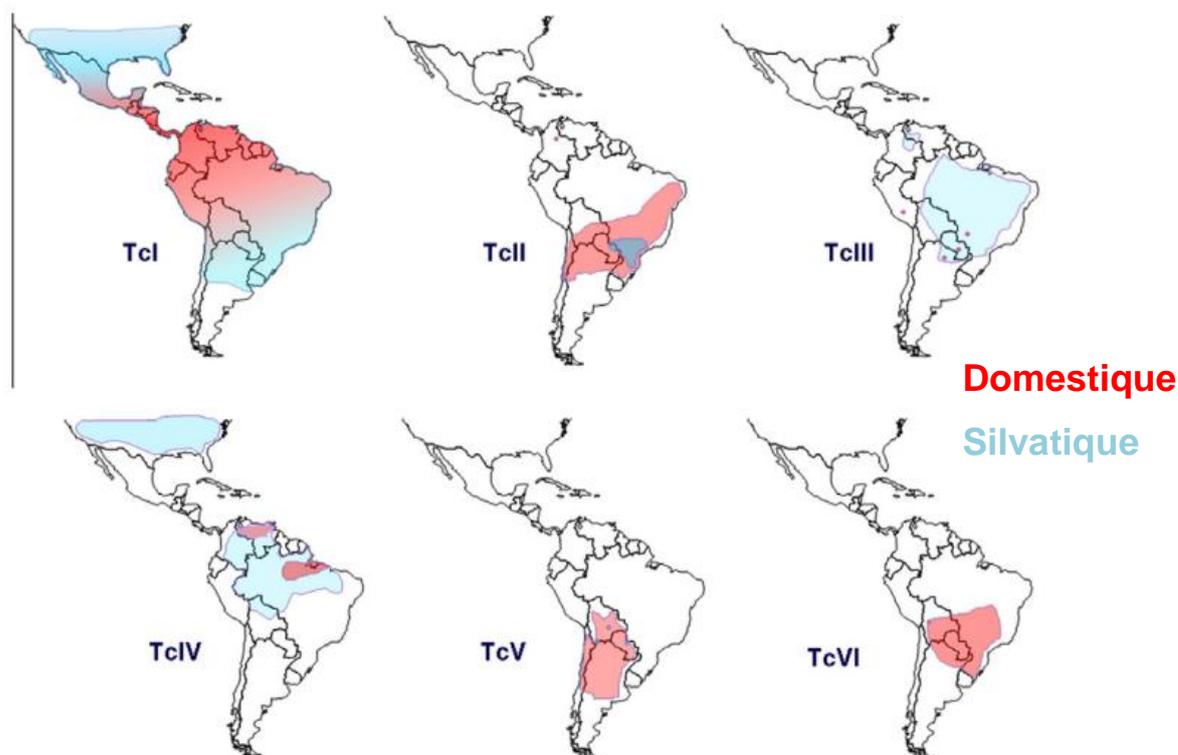


Figure 5. Distribution des différentes souches de *T. cruzi* en Amérique Latine, en fonction de leur reproduction domestique ou sylvatique.

Adaptée de Zingales B. et al., 2011 (22).

La diversité de ces souches se retrouve dans la réponse des organismes à l'infection par *T. cruzi*. En effet, TcIV et TcV sont celles entraînant le moins de complications chez les individus atteints de la maladie de Chagas, alors que TcI et TcII sont les plus virulentes. Ces deux dernières présentent malgré tout des différences : TcI est rarement responsable de formes chroniques digestives, alors que TcII est trouvée chez des individus asymptomatiques, atteints de cardiomyopathies, ou encore de formes digestives (23). De récentes études ont montré une association entre la virulence des souches et la régulation de molécules associées à la réponse immunitaire (24) ou au contrôle du parasitisme (25), mais des études supplémentaires sont nécessaires pour comprendre et caractériser ces différentes réponses.

1.4. Épidémiologie

Les recherches sur la maladie de Chagas se sont étendues après sa découverte au Brésil, en 1909 (1). Ainsi, de 1913 à 1960, elle a été répertoriée dans 13 autres pays d'Amérique Latine ([tableau 1](#)).

| Pays | Première identification d'un cas |
|-------------|----------------------------------|
| Brésil | 1909 |
| El Salvador | 1913 |
| Venezuela | 1919 |
| Pérou | 1919 |
| Costa Rica | 1922 |
| Paraguay | 1924 |
| Guatemala | 1933 |
| Chili | 1937 |
| Mexique | 1938 |
| Bolivie | 1942 |
| Colombie | 1947 |
| Nicaragua | 1949 |
| Argentine | 1949 |
| Honduras | 1960 |

Tableau 1. Chronologie de la première détection de la maladie de Chagas dans divers pays d'Amérique Latine.

Adapté de Lidani KCF et al., 2019 (1).

À partir de 1955, elle a également été identifiée aux États-Unis, d'abord au Texas, puis en Californie (26). Cependant, des traces de *Trypanosoma cruzi* ont été trouvées dans une momie datant des années 800 (entre le Texas et le Mexique), démontrant sa présence bien avant les années 1900 (27). Par ailleurs, les plus anciennes traces de *T. cruzi* datent de 7.750 à 2.000 ans avant JC, entre le Chili, le Pérou et le Brésil. En dehors du continent Américain, le premier cas Européen a été décrit en 1981 (28). Après cette date, le nombre de cas identifiés s'est rapidement accentué, notamment en Espagne, où un grand nombre de personnes en provenance d'Amérique latine, principalement de Bolivie, se sont établies.

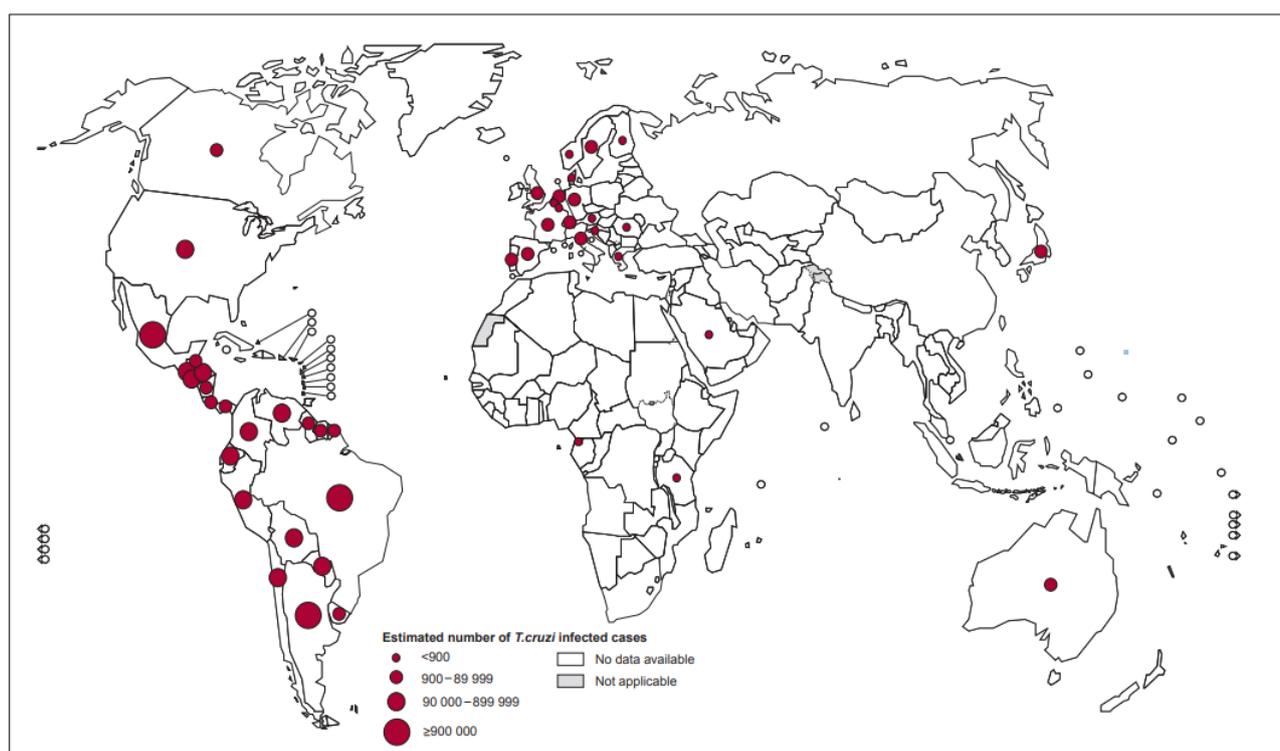


Figure 6. Estimation du nombre de personnes atteintes de la maladie de Chagas à travers le monde en 2018.

Adaptée de Montalvo-Ocototxtle IG et al., 2022 (29).

À l'heure actuelle, la maladie de Chagas s'étend à travers le monde (figure 6). D'après l'Organisation Mondiale pour la Santé, il y a environ 6 à 7 millions de personnes infectées dans le monde, et 10.000 morts par an (WHO, 14 avril 2022). Il est important de noter qu'au-delà du nombre de décès, la qualité de vie des individus atteints de la maladie de Chagas est souvent réduite pendant de nombreuses années.

Le mode d'évaluation DALY (Disability-Adjusted Life Years) est une mesure évaluant l'espérance de vie en bonne santé, soustrayant l'espérance de vie au nombre d'années « perdues » à cause d'une pathologie. Des études ont montré que la maladie de Chagas représentait en moyenne 275.380 DALY en 2019 pour 100.000 habitants à travers le monde, contre 360.870 en 1990 (30), montrant une évolution positive dans la prévention et/ou le traitement de la maladie.

La majorité des cas se situe en Amérique Latine, dans la zone endémique du parasite ($n > 5.7$ millions) (31), notamment en Argentine ($n > 1.5$ millions) et au Brésil ($n > 1.1$ millions). Tout d'abord associés aux populations rurales sud-américaines, les mouvements de populations locaux et internationaux ont amené cette maladie à se développer dans des territoires où elle était absente (32). À l'heure actuelle, environ deux tiers des individus atteints de la maladie de Chagas vivent dans des zones urbaines (14). Au Brésil, la valeur DALY avoisine les 70.7 en 2016 (pour 100.000 habitants), et était de 233.2 en 1990. L'évolution de cette pathologie est fulgurante, puisque le nombre d'années perdues à cause d'un décès est de 58.4 ans en moyenne, et le nombre d'années vécues avec une qualité de vie réduite est de 12.3 (33). Cette distribution est hétérogène à travers le Brésil, les états les plus touchés étant Minas Gerais, São Paulo et Goiás, soit le centre/est du Brésil (**figure 7**).

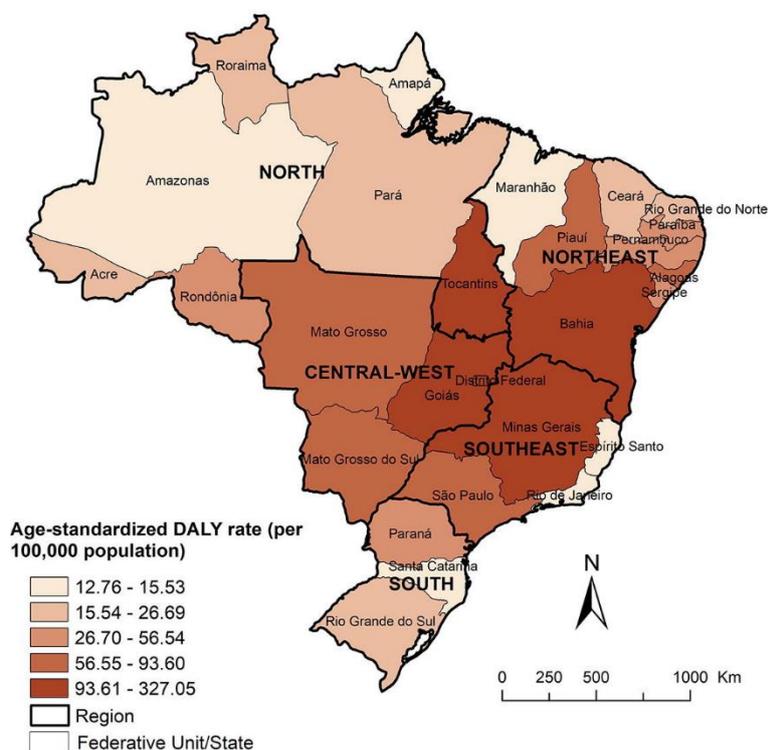


Figure 7. Distribution de la valeur DALY (pour 100.000 habitants) à travers le Brésil en 2016.

Adaptée de Martins-Melo FR et al., 2019 (33).

Aux Etats-Unis, plus de 280.000 Américains provenant d’Amérique Latine sont infectés. 10.000 d’entre eux pourraient avoir été contaminés localement, par transfusion de sang contaminé (34). Toutefois, plusieurs espèces de triatomés ont été recensées dans le pays, et 10 d’entre elles ont montré une infection à *T. cruzi*. M.K.Lynn et al. ont rapporté 76 cas sont probablement infectés directement dans les États du sud, notamment au Texas (48 cas suspectés, 26 confirmés) (35).

En dehors du continent Américain, à l’exception des individus contaminés directement dans la zone endémique du parasite (migration, tourisme...), les personnes atteintes de la maladie de Chagas ont toutes été contaminées de manière non vectorielle. Différentes études estiment environ 100.000 cas de Chagas en Europe (36), notamment en Espagne (entre 12.500 et 25.700 cas), en Italie (entre 4.300 et 4.600 cas) et en France (entre 1.300 et 1.700 cas) (figure 8). Ce sont également les pays qui accueillent le plus de migrants sud-américains (37).

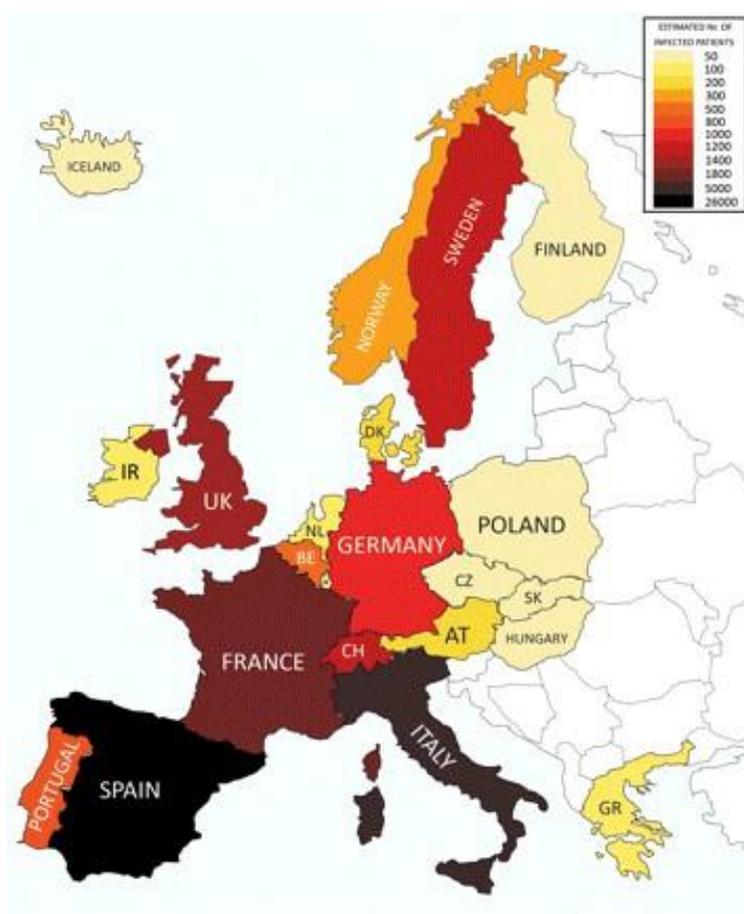


Figure 8. Estimation du nombre d’individus atteints de la maladie de Chagas en Europe.

Adaptée de Roberto A. et al., 2008 (37).

Au Japon, environ 3.000 personnes sont contaminées par *T. cruzi* (38). En Australie et en Nouvelle-Zélande, 116.000 et 6.300 individus nés en Amérique Latine respectivement sont contaminés par *T. cruzi*. L'absence de diagnostic systématique des migrants provenant de régions endémiques du parasite, de dépistage lors d'une grossesse et/ou de contrôles des banques de sang tendent à sous-estimer le nombre de cas de Chagas.

De plus, le poids économique que représentent les soins des patients atteints de la maladie de Chagas est non négligeable. En effet, le coût des soins au cours de la vie d'une personne infectée par *T. cruzi* équivalent à 2.600\$ en Amérique Latine, 14.950\$ en Europe ou 19.950\$ aux Etats-Unis, Canada et Australie. À cela s'ajoute la perte financière associée à la détérioration de la qualité de vie de ces patients, ce qui fait grimper ces coûts. Au total, cette pathologie coûte, au cours de la vie d'un individu, 24.200\$ en Amérique Latine, 77.000\$ en Europe et 91.500\$ aux États-Unis, Canada et Australie. Pour le Brésil, cela se traduit par un fardeau économique annuel de 1.7 milliards de dollars, et pour les États-Unis, 862 millions de dollars. À l'échelle mondiale, la charge économique s'élève à 188,8 milliards de dollars par an, dépassant de loin celle d'autres maladies infectieuses telles que le choléra et le rotavirus, et rivalisant avec certains cancers (comme le cancer de l'utérus et de la bouche) (39).

Qu'il s'agisse de zones endémiques ou non endémiques, l'impact sur la santé et l'économie de la maladie de Chagas demeure significatif. L'amélioration des mesures de contrôle sanitaire pourrait non seulement contribuer à améliorer la qualité de vie des individus infectés, mais aussi à alléger le fardeau économique associé à cette maladie.

Pour conclure, la maladie de Chagas est une maladie parasitaire due à un protozoaire, *Trypanosoma cruzi*. Elle est principalement transmise à travers un vecteur en zone endémique, le réduve, mais peut également être contractée par l'alimentation ou du sang et/ou organes contaminés. Les différents mouvements de population et, plus récemment, le réchauffement climatique ont permis à cette maladie de se propager à travers le monde. Le chapitre suivant permettra d'introduire les différents aspects médicaux relatifs à cette pathologie, des symptômes aux traitements disponibles.

Chapitre 2 : L'évolution de la maladie

Le chapitre précédent a introduit la maladie de Chagas, de sa découverte à sa répartition dans le monde. Cette deuxième partie se concentrera sur les différentes formes de la maladie, les symptômes en résultant et les différents traitements disponibles à l'heure actuelle.

2.1. La phase aiguë

La phase aiguë apparaît directement après l'infection, lorsque le parasite entre dans l'organisme de l'hôte. La période d'incubation s'étend sur 2 à 3 semaines en fonction du mode d'infection (congénital, vectoriel ou oral), mais peut aller jusqu'à 160 jours en cas de transmission sanguine (40). Le plus fréquemment, cette phase dure 4 à 8 semaines et est majoritairement asymptomatique (60-70% des individus infectés) (41). Les symptômes les plus courants sont communs à toutes les infections, comme de la fièvre, de la fatigue, une adénopathie (augmentation du volume du nœud lymphatique), parfois une hépatomégalie (hypertrophie du foie), une splénomégalie (augmentation du volume de la rate), une dyspnée, une myalgie et des nausées (transmission orale). Dans les cas de contamination vectorielle, au niveau du site d'infection de *T. cruzi*, un chagome peut apparaître (**figure 9A**). Il s'agit d'une lésion cutanée qui ressemble à un furoncle. Si le parasite est entré au niveau de la conjonctive, le signe de Romaña apparaît, caractérisé par une inflammation de la conjonctive formant un œdème (40) (**figure 9B**).

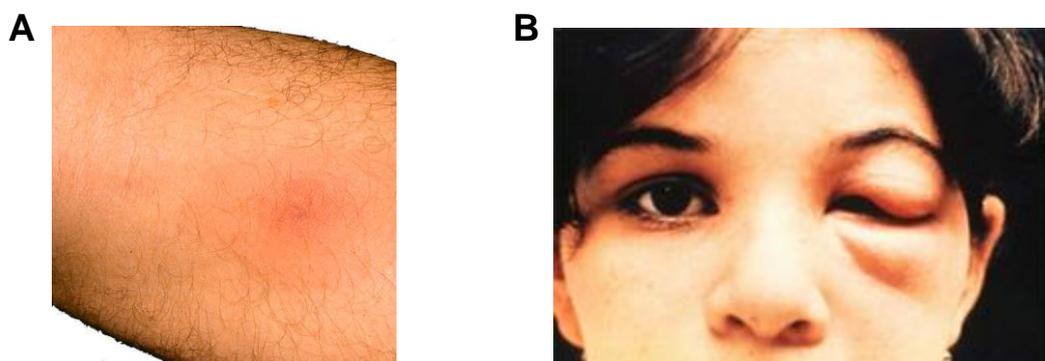


Figure 9. Photographies montrant un chagome (A) et le signe de Romaña (B).

Adaptée de Rossi MA et al., 2003 (42).

Dans de rares cas plus sévères (5%), une myocardite, une méningo-encéphalite, une anémie, une thrombopénie ou une hyperlymphocytose peuvent survenir. En cas de contamination orale, des douleurs thoraciques, abdominales ou une hémorragie digestive peuvent apparaître. Le risque de décès est relativement faible en phase aiguë (0.2 à 0.5%), mais augmente en cas de contamination orale (1 à 35%) (40).

2.2. La phase chronique

La phase chronique, succédant la phase aiguë, peut être symptomatique ou asymptomatique. La majorité (70%) des individus infectés ne développe aucun symptôme tout au long de leur vie. Cependant, environ 30% d'entre eux développeront des manifestations cliniques, le plus souvent cardiaques (20-30%), mais également digestives (5-20%), cardio-digestives ou neurologiques, qui apparaissent des années, voire des décennies après l'infection (43,44).

2.2.1. Les complications cardiaques : la Cardiomyopathie Chagastique Chronique

Le cœur est un muscle permettant de faire circuler le sang dans l'organisme. Il est divisé en 2 parties, droite et gauche, chacune composée de deux cavités : l'oreillette et le ventricule. Le sang pauvre en oxygène arrive dans l'oreillette puis le ventricule droit, pour être envoyé vers les poumons et être approvisionné en oxygène. Le sang oxygéné arrive alors dans l'oreillette puis le ventricule gauche, pour être envoyé dans les artères. Entre chaque chambre, des valves s'ouvrent et se ferment pour éviter un retour en arrière du sang. Le mouvement sanguin se fait grâce aux contractions du muscle cardiaque, appelé le myocarde. Les cellules le composant sont essentiellement des cardiomyocytes, des cellules contractiles autonomes.

La Cardiomyopathie Chagastique Chronique (CCC) est une cardiomyopathie dilatée et inflammatoire. Elle est caractérisée par une dilatation du muscle myocarde, réduisant la capacité du cœur à propulser le sang dans l'organisme. Des événements thromboemboliques sont assez fréquents, soit l'apparition de caillots dans des veines (thrombose veineuse) ou dans des artères pulmonaires (embolie pulmonaire). Une destruction des cellules neuronales cardiaques est observée (45), ainsi que des problèmes d'arythmie, dus à des anomalies du système de conduction (46).

Dans les cas plus sévères, une hypertrophie des 4 cavités cardiaques est constatée (oreillettes et ventricules) avec une nécrose des cardiomyocytes. Cela conduit au développement de fibroses (tissu se développant dans les zones où les cellules détruites ne se régénèrent pas). L'atteinte du cœur peut conduire à une insuffisance cardiaque, et la mort subite.

Par ailleurs, dans le tissu cardiaque apparaissent à la fois des foyers de formes amastigotes de *T. cruzi* et un infiltrat de cellules immunitaires, comprenant notamment des lymphocytes et des macrophages (**figure 10**), induisant une réaction inflammatoire continue. L'intensité de la réponse inflammatoire semble par ailleurs corrélée avec la concentration de parasites (45).

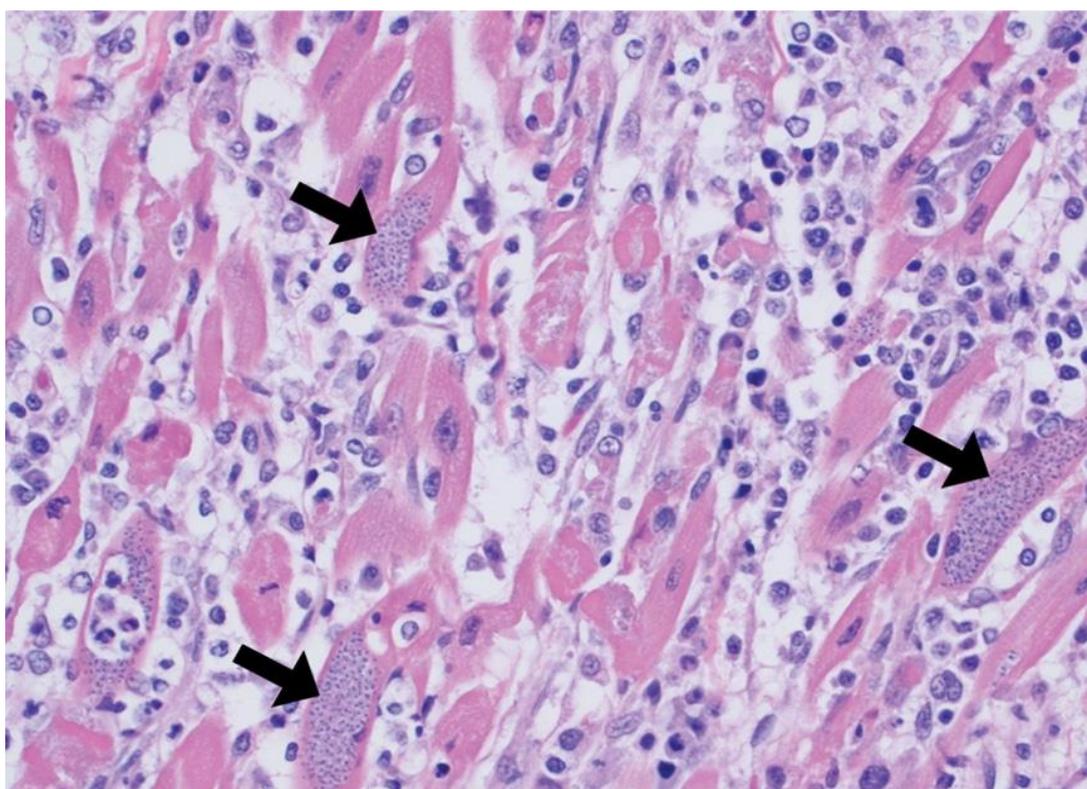


Figure 10. Échantillon histologique du ventricule droit d'un chien infecté par *T. cruzi*.
*Les cellules roses sont des cardiomyocytes, et les violettes des cellules immunitaires. Les flèches
montrent les foyers amastigotes de T. cruzi dans les cardiomyocytes.*
Adaptée Saunders AB et Hamer SA., 2020 (47).

Les premiers symptômes sont de la fatigue, des syncopes (pertes de connaissance), des palpitations, des vertiges et un accident vasculaire cérébral (AVC). La détérioration du cœur est généralement progressive, et finit par induire des douleurs thoraciques anormales, un essoufflement (dyspnée), des œdèmes, et une insuffisance cardiaque congestive (40).

Il s'agit d'une maladie grave, conduisant le plus souvent au décès. Le taux de survie des patients atteints de CCC est plus faible que celui de patients atteints d'autres pathologies cardiaques (**figure 11**).

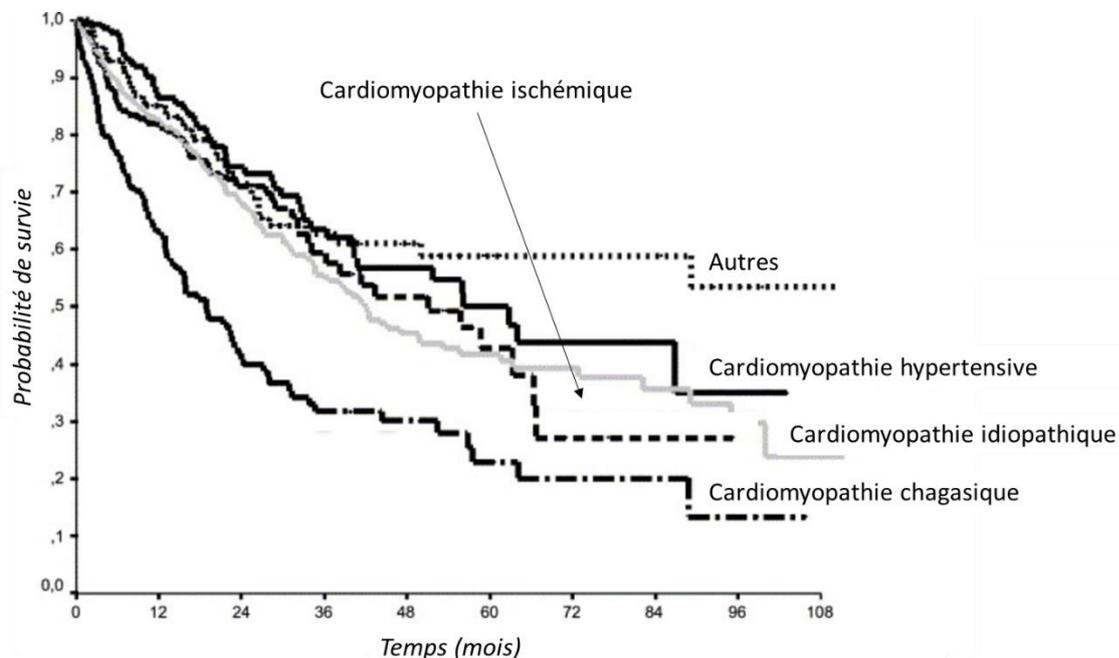


Figure 11. Courbe de survie temporelle des patients atteints de CCC et d'autres pathologies cardiaques. Adaptée de Freitas HFG et al., 2005 (48).

2.2.2. Les complications gastro-intestinales

Le tube digestif est un ensemble de 7 organes ayant pour but d'ingérer, digérer, absorber et éliminer les aliments. Il comprend la bouche, le pharynx, l'œsophage, l'estomac, l'intestin grêle, le gros intestin et le rectum. L'activité motrice est contrôlée par le système nerveux entérique, composé de deux plexus ganglionnaires (rassemblement de nerfs et vaisseaux), le plexus myentérique et le plexus sous-muqueux. Le premier se trouve entre les muscles, et le deuxième sous la muqueuse. La majorité du système gastro-intestinal peut être affectée par la maladie de Chagas, bien que l'œsophage (méga-œsophage) et le côlon (mégacôlon) soient majoritairement touchés. Les principaux symptômes sont des nausées, vomissements, régurgitations, perte de poids et constipations chroniques. Les organes touchés sont hypertrophiés, avec des troubles des muscles moteurs entraînant des altérations de la motilité (49). Ces troubles musculaires sont dus à la destruction de neurones, conduisant à une dénervation des plexus myentériques et sous-muqueux (fibres nerveuses dans les tissus musculaires situés le long de la paroi du tube digestif).

Un double effet parasitaire et immunitaire semble être à l'origine de cette destruction neuronale (44). Cette atteinte n'est pas associée à des taux de mortalité plus élevés, mais diminue beaucoup la qualité de vie.

2.3. Diagnostic

Comme discuté précédemment, plus de 30% des individus infectés développent les formes chroniques symptomatiques, pouvant mener au décès, des années après l'infection. Il est donc primordial que le diagnostic soit le plus précoce possible. Nous verrons ici les différentes méthodes de diagnostic de la maladie de Chagas.

2.3.1. Evaluation de la contamination par *T. cruzi*

Il y a plusieurs méthodes permettant d'identifier une infection par *T. cruzi*, leur efficacité différant en fonction du stade de la maladie (50), la parasitémie étant très élevée en phase aiguë, mais relativement faible, voire nulle, en phase chronique. Nous détaillerons ici les 3 méthodes les plus utilisées : la microscopie, la sérologie et la PCR.

- Détection par microscopie

En phase aiguë, le diagnostic peut être directement fait par microscopie. Un prélèvement sanguin est coloré au giemsa et déposé sur une lamelle. La présence de *T. cruzi* est directement observable au microscope (**figure 12**).

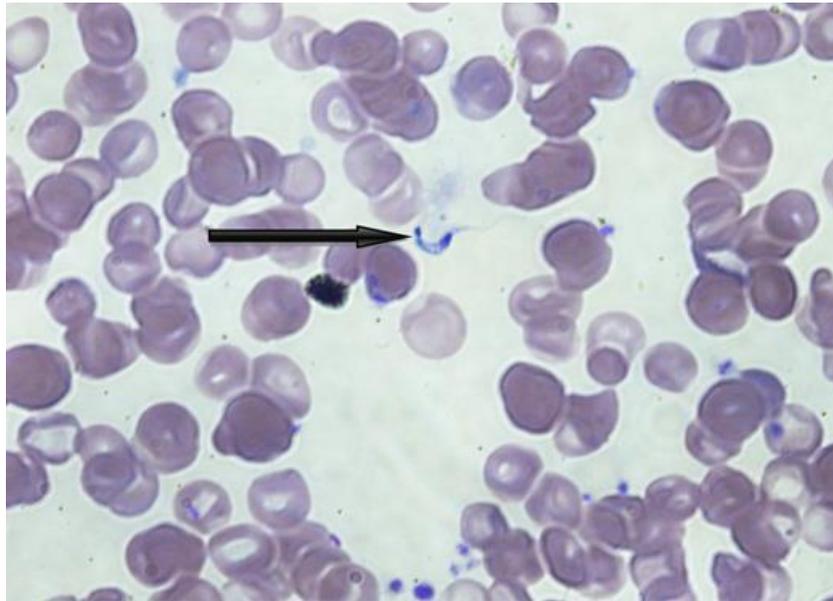


Figure 12. Photographie de la forme trypanomastigote de *T. cruzi* présent dans un échantillon sanguin coloré au giemsa.

Adaptée de Alarcón A. et al., 2016 (51).

- Détection par sérologie

La sérologie est une analyse visant à détecter dans des échantillons sanguins des anticorps spécifiques d'un pathogène. Elle est utilisée pour détecter une potentielle contamination en phase chronique. Les tests les plus souvent utilisés sont les tests ELISA (Enzyme-Linked Immunosorbent Assay), IFAT (Indirect Fluorescent Antibody Test) et HA (Hemagglutination Assay). Les deux premiers consistent à fixer des antigènes de *T. cruzi* sur une surface solide, puis à ajouter des échantillons sanguins de patients. Si les patients sont contaminés, des anticorps anti-*T. cruzi* se lieront aux antigènes fixés. L'ajout d'un anticorps secondaire couplé à une enzyme va permettre de détecter les complexes anticorps-antigènes. La réaction chimique en résultant va produire un signal colorimétrique (ou fluorescence) mesurable par un spectrophotomètre. Le test HA, en revanche, fixe les antigènes à des globules rouges, qui sont mélangés à des échantillons sanguins de patients. S'il y a liaison anticorps-antigène, les globules rouges vont s'agglutiner. Ce test est plus rapide que les deux premiers, et aussi fiable qu'un test IFAT. Cependant, le test ELISA est le plus précis (spécificité de 99% contre 78%), ce qui explique qu'il soit le plus utilisé (52). La PAHO (Pan American Health Organization) recommande de combiner deux tests sérologiques différents.

S'il y a un conflit dans les résultats, un troisième test doit être effectué (50). Le test sérologique présente des limites dans la détection de la transmission congénitale, du fait de la présence des anticorps anti-*T. cruzi* dans le sang maternel (53).

- Détection par PCR

En phase aiguë, la détection de *T. cruzi* peut se faire par PCR (Polymerase Chain Reaction). Le principe est d'amplifier un segment particulier de l'ADN du parasite sur un échantillon tissulaire ou sanguin. Cette méthode est plus précise que la microscopie, avec une précision approchant les 100% selon les kits, mais également plus coûteuse. Cette méthode est également utilisée dans d'autres situations : elle permet de contrôler l'efficacité d'un traitement antiparasitaire, ou de détecter une potentielle transmission congénitale.

2.3.2. Diagnostic des cardiomyopathies chagasiennes chroniques

Les premiers signes d'une cardiomyopathie chagasique chronique sont observables à l'électrocardiogramme (ECG). Il s'agit d'un examen évaluant l'activité électrique du cœur au fil des battements. Cette activité est mesurée par le biais d'électrodes placées sur le torse et les membres. L'ECG est représenté par une courbe illustrant l'activité électrique en millivolts (axe des ordonnées) en fonction du temps en secondes (axe des abscisses). En fonction des axes créés par le positionnement des différentes électrodes (ex. bras droit et bras gauche, bras droit et jambe gauche...), plusieurs mesures sont prises, donnant différentes courbes complémentaires. Globalement, on observe sur l'ECG le complexe QRS, indicateur du bon fonctionnement du cycle cardiaque. Plus précisément, ce complexe enregistre l'activité électrique des ventricules cardiaques pendant leur contraction. L'onde P précède ce complexe et illustre la contraction des oreillettes, tandis que l'onde T suit le complexe QRS et indique le relâchement des ventricules ([figure 13](#)).

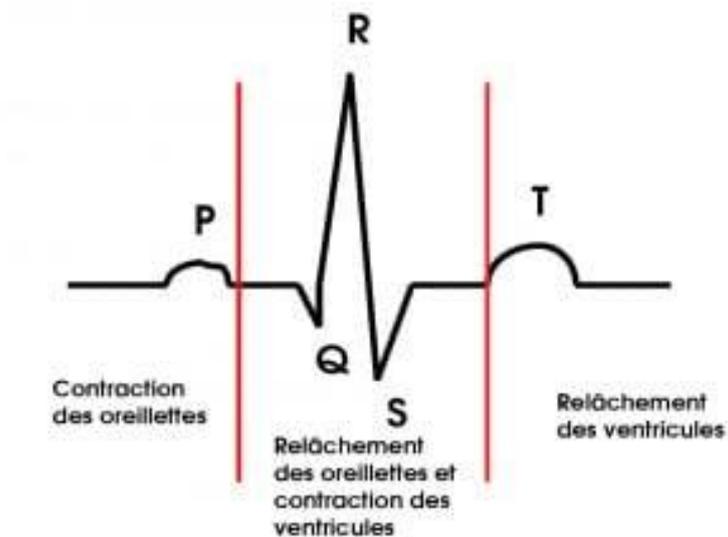


Figure 13. Représentation schématique du complexe QRS observé sur un ECG normal.
Source : Mediprostore.

Dans le cas d'une CCC, le voltage QRS est bas et l'onde T a plus d'amplitude (45) (figure 14). Chez des personnes atteintes de la maladie de Chagas, ce signal est très important, car il est associé à un risque élevé de développer une cardiomyopathie. Le diagnostic des CCC avec un ECG est fréquemment utilisé en examen préliminaire dû à son faible coût et sa facilité d'utilisation et d'interprétation. Toutefois, cet examen n'apporte pas d'informations détaillées sur l'avancement de la maladie et n'est pas toujours précis. Par exemple, Rojas LZ et al. ont réalisé une méta-analyse montrant qu'en moyenne, la valeur prédictive positive de l'ECG pour les CCC (nombre de vrais positifs divisés par la somme des vrais et faux positifs) est d'environ 80% (54). Il est donc nécessaire d'effectuer des examens complémentaires.

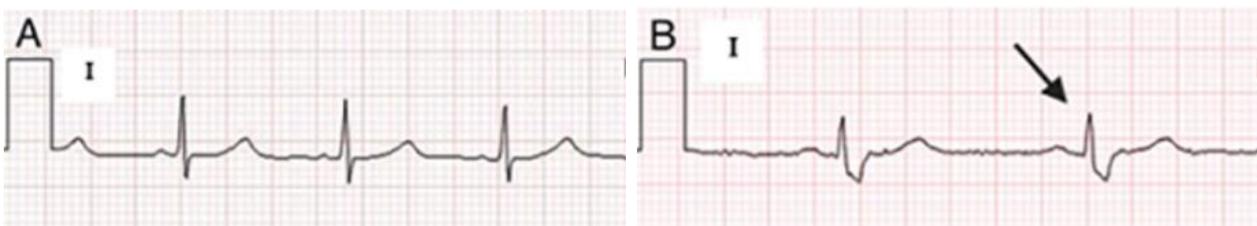


Figure 14. Comparaison d'un ECG normal (A) avec un ECG de patient atteint de CCC (B). La diminution du voltage QRS et l'amplitude de l'onde T sont visibles.

Adaptée de Traina MI et al., 2017 (55).

Parmi ces examens, l'échographie est indiquée pour tous les patients atteints de la maladie de Chagas, mais n'est pas systématique chez ceux connus pour être atteints de CCC. Des anomalies peuvent être visibles dans la paroi du ventricule gauche, même chez des patients ayant un ECG normal. L'échographie permet également d'observer une dilatation cardiaque dans des stades plus avancés. Une IRM peut également être réalisée. L'avantage de l'IRM est qu'il permet d'avoir une analyse de haute qualité de l'anatomie et la fonction cardiaque. Il est donc possible de déterminer le volume de la cavité cardiaque, le degré de contractilité, ainsi que les zones du cœur présentant une fibrose (56).

Plusieurs métriques existent pour classer l'évolution des CCC. Par exemple, la fraction d'éjection cardiaque du ventricule gauche (Left Ventricular Ejection Fraction, LVEF), soit le niveau de sang éjecté par ce ventricule, reflète la bonne fonction contractile du cœur. Différentes études classent les CCC sous forme modérée lorsque la LVEF est supérieure ou égale à 40%, ou sévère lorsqu'elle est inférieure à ce seuil (57). Un autre classement, plus précis, s'appuie sur plusieurs paramètres cliniques pour évaluer le niveau d'atteinte cardiaque, comme les changements au niveau de l'ECG, la fraction d'éjection cardiaque, et les observations faites à l'échographie (45). Enfin, la classification NYHA (New York Heart Association) est fréquemment employée pour évaluer le niveau d'insuffisance cardiaque, chez les CCC mais également pour d'autres atteintes cardiaques. Cette dernière prend également en compte la sévérité des symptômes ressentis par le patient.

2.3.2. Diagnostic des atteintes digestives chroniques

Le diagnostic des atteintes digestives est essentiellement réalisé par radiographie et endoscopie, afin d'identifier une potentielle hypertrophie des organes, et évaluer leur état. Dans le cas d'un mégaoesophage, une manométrie est employée pour évaluer la contraction de l'œsophage (58).

2.3.3. L'utilisation de biomarqueurs spécifiques des différentes formes de la maladie

Les différentes méthodes de diagnostic présentées précédemment, notamment pour les formes chroniques, nécessitent à la fois des moyens humains et matériels conséquents.

Dans les zones endémiques, la majorité des individus atteints de la maladie de Chagas vivent en zone rurale, et n'ont qu'un accès limité aux services de santé (59). La découverte et l'utilisation de biomarqueurs spécifiques des différentes formes de la maladie ainsi que de son évolution seraient un atout dans le diagnostic de ces populations. Cela permettrait également un dépistage simplifié et moins coûteux que les techniques actuelles.

Dans le cas des CCC, de nombreux biomarqueurs ont déjà été proposés dans la littérature. Ribeiro ALP et al. ont étudié l'utilisation du BNP (Brain Natriuretic Peptide), un marqueur bien connu des pathologies cardiaques (60), comme marqueur des dysfonctions du ventricule gauche (61). Avec cette molécule, ils ont pu prédire le développement de CCC une sensibilité de 80% et une spécificité de 97%. Plus tard, Keating SM et al. ont confirmé l'intérêt de plusieurs biomarqueurs cardiovasculaires dans la prédiction des CCC, tels que la créatine phosphokinase MB (CK-MB), la troponine, la myoglobine, la protéine d'adhérence des cellules vasculaires 1 (VCAM) et surtout le NT-proBNP, un fragment du BNP (62). Toutefois, l'ensemble de ces biomarqueurs démontrent d'une affection cardiaque, mais ne sont pas spécifiques aux CCC (63,64).

Différentes protéines inflammatoires ont été associées aux formes cardiaques, comme l'interféron gamma (IFN- γ), le tumor necrosis factor alpha (TNF- α), l'interleukine 6 (IL-6) et l'interleukine 10 (IL-10) (62,65). Le niveau d'IFN- γ a été également inversement corrélé à la fraction d'éjection cardiaque (65), montrant un intérêt dans l'étude de l'évolution des CCC. La métalloprotéinase 9 (MMP-9) a également été proposée comme marqueur des CCC sévères, alors que la métalloprotéinase 2 (MMP-2) semble caractéristique des formes chroniques asymptomatiques ou CCC précoces (66,67). En dehors des protéines, plusieurs micro-ARNs (ARN non codants), incluant miR-1, miR-208b, et miR-133a-2, ont également été associés aux CCC, comparés à des contrôles sains et des cardiomyopathies dilatées non inflammatoires. Ces derniers pourraient dans le futur devenir des marqueurs des formes cardiaques de la maladie de Chagas (68). Pris communément, ces potentiels biomarqueurs pourraient à la fois discriminer les CCC des autres formes chroniques, mais également des cardiomyopathies dilatées non inflammatoires. Des études supplémentaires restent nécessaires pour confirmer leur spécificité aux CCC.

2.4. Traitements

Il n'existe, à l'heure actuelle, aucun vaccin permettant de lutter contre les formes graves de la maladie de Chagas. La meilleure méthode pour lutter contre cette pathologie est l'élimination du parasite.

2.4.1. Les traitements antiparasitaires

Il n'existe que deux traitements antiparasitaires efficaces dans le cadre de la maladie de Chagas : le benznidazole et le nifurtimox. Le benznidazole est un promédicament, cela signifie qu'il est ingéré sous forme inactive, et s'active en utilisant des molécules présentes dans l'organisme. Dans ce cas précis, l'élément activateur est une protéine présente chez *T. cruzi*. Le benznidazole va entraîner la destruction du parasite sous toutes ses formes par le biais de dommages oxydatifs. En revanche, il entraîne parfois des effets secondaires, notamment des dermatites allergiques (20-30% des cas), des polyneuropathies (10%) et une neutropénie (baisse du nombre de globule blanc). Le nifurtimox va participer à la génération de radicaux libres, qui vont détruire le parasite. Ce médicament est plus toxique que le benznidazole, ce qui a entraîné la fin de sa commercialisation dans plusieurs pays d'Amérique latine dans les années 1980. Cependant, faute d'alternative, son utilisation est toujours préconisée. Il entraîne dans 50% des cas des troubles gastrointestinaux (anorexie, perte de poids, nausées, vomissements, douleurs abdominales), ainsi que des insomnies, de l'irritabilité et de la désorientation (69,70).

L'efficacité de ces deux médicaments est variable. Globalement, le benznidazole est plus efficace, avec un taux de guérison allant de 60 à 94% chez les enfants et de 2 à 40% chez les adultes chroniquement infectés. Le nifurtimox, quant à lui, a une efficacité de 86% chez les enfants, et de 7-8% chez les adultes (69). Malheureusement, dû à leur coût et aux difficultés d'approvisionnements, ils ne sont pas disponibles dans toute l'Amérique Latine. On estime que seulement 1% de la population infectée par *T. cruzi* est traitée (71). Comme mentionné par Farmer P., "les frontières politiques sont des membranes semi-perméables, souvent très ouvertes aux maladies et pourtant fermées à la libre circulation des traitements" (32).

2.4.2. Les traitements des atteintes cardiaques

Il n'existe à l'heure actuelle aucun traitement permettant de guérir les CCC. L'usage des traitements antiparasitaires n'a que peu d'effets, les dommages tissulaires étant majoritairement dus à une forte réponse inflammatoire (72). Toutefois, Hasslocher-Moreno AM. et al. ont récemment suggéré que la prise de benznidazole en phase chronique asymptomatique tend à limiter le risque de développer des CCC (73). Par ailleurs, plusieurs thérapies sont proposées pour limiter l'évolution de la maladie et lutter contre les différents symptômes. Ces approches sont les mêmes que celles appliquées dans le traitement de l'insuffisance cardiaque dues à d'autres pathologies.

Pour les patients atteints d'arythmie et d'insuffisance cardiaque, l'utilisation de bêta-bloquants, des médicaments bloquant l'action des neurotransmetteurs et régulant ainsi le rythme cardiaque, est souvent prescrite. Il a été démontré que l'usage de ces derniers améliore le taux de survie des patients atteints de CCC, le rendant comparable au taux de survie de patients atteints d'autres pathologies cardiaques (74). En parallèle des bêta-bloquants, des inhibiteurs de l'enzyme de conversion de l'angiotensine sont également donnés. Ils entraînent une dilatation des vaisseaux sanguins, et réduisent donc la force nécessaire au cœur pour faire circuler le sang efficacement (56).

Dans le cadre d'insuffisances cardiaques autres que des CCC, deux molécules, le sacubitril et le valsartan, sont généralement prescrites. Récemment, une étude réalisée sur 17 patients atteints de CCC a montré que leur utilisation améliore la fraction d'éjection cardiaque, et donc le pronostic de ces individus (75). Toutefois, leur utilisation dans le cadre des CCC n'a pas encore fait l'objet d'essais cliniques appropriés.

L'implantation de pacemakers ou de défibrillateurs est également employée pour aider les patients qui présentent un trouble du rythme anormalement lent, et lutter contre l'arrêt cardiaque. Pour les patients ayant une CCC très avancée, la transplantation cardiaque est le seul recours (46), avec une survie de 71% au bout d'un an, et de 46% après 10 ans. De nombreux cas de réactivations ont été observés, marqués par la détection de nouvelles niches amastigotes dans le cœur greffé, initialement sain. Une analyse portant sur 107 patients ayant subi une transplantation cardiaque et présentant une CCC a révélé que 43 d'entre eux (40%) ont connu une réactivation de la pathologie (76).

2.4.3. Les traitements des atteintes digestives

Tout comme pour les formes cardiaques, il n'existe pas de traitements permettant de guérir ces atteintes. Dans le cadre d'atteintes œsophagiennes, des nitrates ou de la nifédipine peuvent être prescrits pour réduire la pression de l'œsophage. Le plus fréquemment, une opération par endoscopie est réalisée. Il peut s'agir d'une dilatation œsophagienne par ballon, consistant à introduire de petits ballons dans le conduit œsophagien et à les gonfler pour détendre les muscles. Une myotomie de Heller peut également être réalisée, soit une légère dissection des muscles permettant de les détendre et de faciliter le passage des aliments. En ce qui concerne le mégacôlon, des laxatifs sont prescrits pour faciliter le transit. Dans les cas les plus graves, une opération chirurgicale est nécessaire (77).

Chapitre 3. De la défense de l'hôte au développement des cardiomyopathies chagasiques chroniques

Les chapitres précédents ont introduit la maladie de Chagas, son agent pathogène et les symptômes apparaissant en phase chronique. Cette partie se concentrera sur les mécanismes moléculaires et cellulaires de l'hôte faisant suite à l'infection par *T. cruzi*, allant de la phase aiguë jusqu'au développement des formes symptomatiques, et plus particulièrement des CCC.

3.1. L'infection de l'hôte par le parasite et le déclenchement de la réponse immunitaire innée

3.1.1. L'entrée du parasite dans les cellules

Le processus d'entrée de *T. cruzi* dans les cellules est complexe, et implique un large groupe de glycoprotéines et de protéases de surface. Lorsque le parasite entre dans l'organisme, il infecte en premier les cellules à proximité (cellules immunitaires, épithéliales), puis des cellules plus distantes (musculaires, fibroblastes). Deux types d'entrée sont à distinguer : la phagocytose, et la pénétration active dans les cellules.

La phagocytose est un mécanisme de défense immunitaire mis en place par les cellules phagocytaires (neutrophiles, macrophages, cellules dendritiques). Il consiste à internaliser des éléments pathogènes afin de les éliminer. La phagocytose se passe en trois étapes : l'adhésion, l'ingestion, et la digestion. L'adhésion se fait par contact physique entre les motifs moléculaires associés aux pathogènes (PAMP, Pathogen Associated Molecular Patterns) du parasite et les récepteurs de motifs moléculaires (PRR, Pattern Recognition Receptor) des cellules immunitaires. La membrane externe va alors se déformer pour former une vacuole, faisant entrer le parasite dans la cellule (78). La vacuole ainsi formée se nomme le phagosome. Ce dernier se déplace vers les lysosomes, contenant des enzymes digestives. Il s'agit d'hydrolases, qui rompent les liaisons chimiques des macromolécules, incluant des protéines, des lipides et des glucides. La fusion du phagosome et du lysosome, appelée un phagolysosome, conduit à la destruction du pathogène par les enzymes digestives (**figure 15**). Bien que la phagocytose soit une réponse de défense de l'organisme, elle ouvre la voie au parasite pour entrer dans les cellules immunitaires.

Une fois à l'intérieur, *T. cruzi* peut mettre en place un mécanisme de défense antioxydant dont l'efficacité peut empêcher sa destruction, lui permettant ainsi de se multiplier dans le phagolysosome avant de poursuivre son infection dans d'autres cellules.

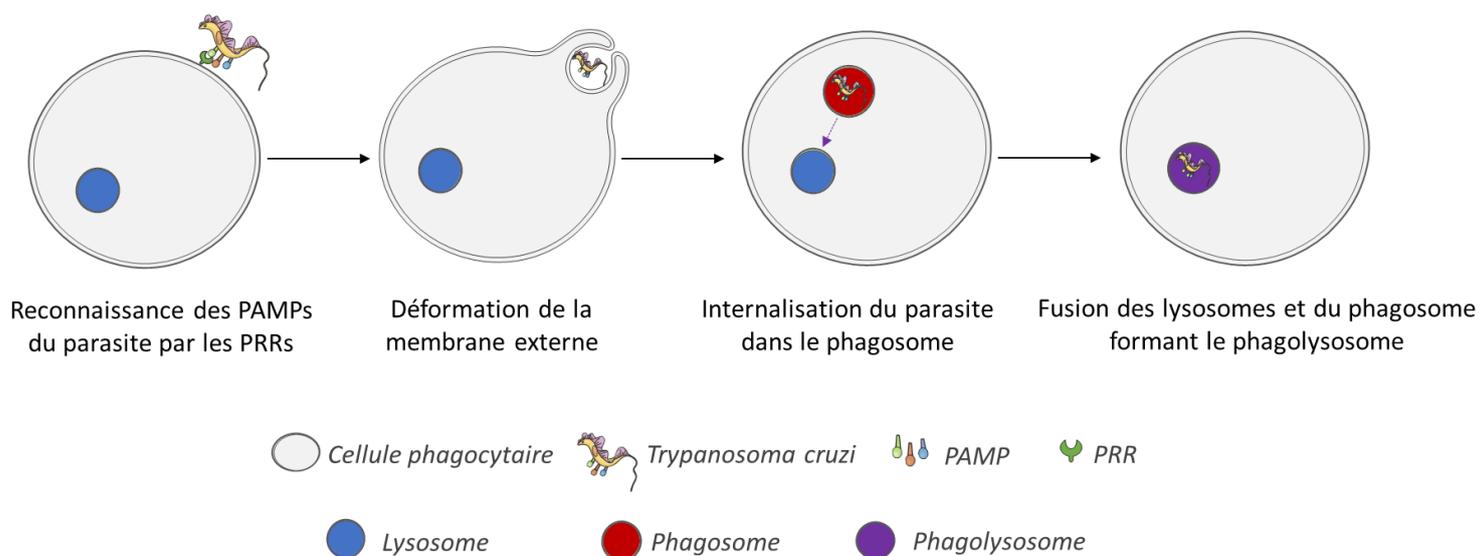


Figure 15. Pénétration de *T. cruzi* dans les cellules par phagocytose.

La pénétration active survient dans les cellules dépourvues de mécanismes de phagocytose, notamment les cellules épithéliales, fibroblastes et musculaires. Ces tissus sont parmi les premiers à être touchés une fois que le parasite a réussi à envahir l'organisme. Le processus d'entrée du parasite n'est pas totalement compris, et plusieurs modèles ont été proposés (79), dépendants ou indépendants des lysosomes. Dans le modèle dépendant, des molécules de surface de *T. cruzi* se fixent sur la surface de la cellule hôte. Cette interaction va conduire à la génération d'IP3 (1,4,5-inositol triphosphate), entraînant la libération de Ca^{2+} dans le cytosol par le réticulum endoplasmique, ainsi qu'au recrutement des lysosomes vers la surface. L'élévation de la concentration du Ca^{2+} cytosolique favorise la rupture du cytosquelette d'actine. Les lysosomes, quant à eux, fournissent une membrane au parasite, lui permettant d'entrer dans la cellule (80) et formant la vacuole parasitophore (**figure 16A**). Dans le modèle indépendant des lysosomes, le parasite utilise directement la membrane de la cellule pour former sa vacuole. La fusion avec les lysosomes se fera dans un second temps, dans le cytosol de la cellule (**figure 16B**).

Quel que soit le mécanisme d'entrée, la fusion avec les lysosomes est essentielle pour le parasite : elle permet de le retenir dans la cellule hôte et de réaliser une réplication efficace.

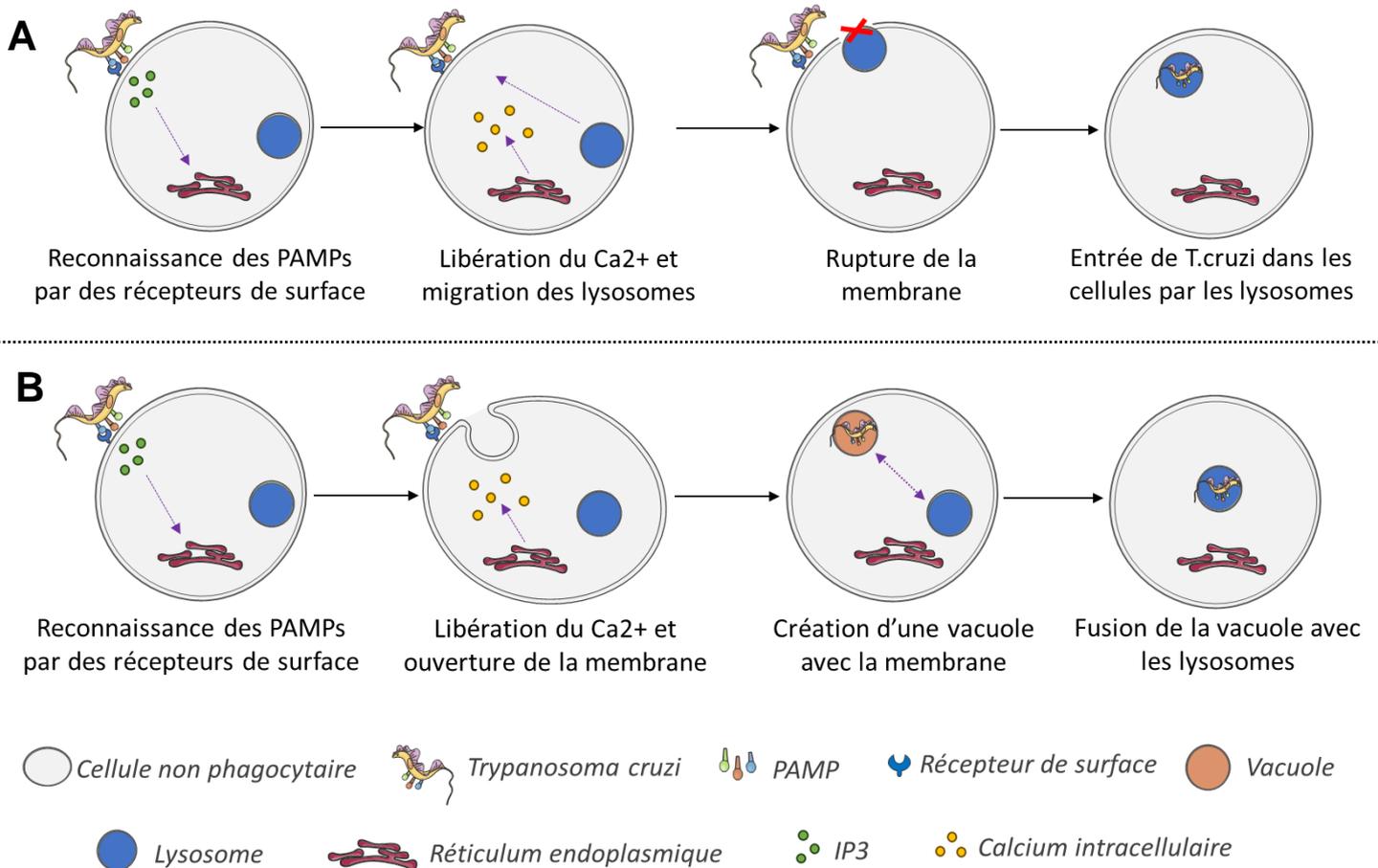


Figure 16. Pénétration active de *T. cruzi* dans des cellules non phagocytaires A) par le biais des lysosomes ou B) en utilisant la membrane cellulaire pour former une vacuole.

3.1.2. Le déclenchement de la réponse immunitaire innée

T. cruzi présente plusieurs protéines de surface qui sont reconnues par les cellules immunitaires, telles que la mucine rattachée par le glycolipide GPI (glycosylphosphatidylinositol) ainsi que le GIPL (glycoinositolphospholipides libre) (81). Ces protéines vont activer des PRR spécifiques, les TLR (Toll Like Receptor). Il s'agit de récepteurs membranaires se situant à la surface des cellules ou dans des compartiments intracellulaires, tels que le réticulum endoplasmique ou le lysosome. Ils sont présents sur les membranes de la majorité des cellules, impliquées dans la réponse immunitaire ou non. Chez l'homme, la famille des TLR comprend 10 récepteurs différents, numérotés de 1 à 10 (82). Ces récepteurs sont sensibles à différents ligands, provoquant une cascade de réaction différente, et apportant ainsi une certaine adaptabilité de la réponse immunitaire innée.

Dans le cadre de l'infection par *T. cruzi*, les GPI vont se fixer sur les TLR 2 et 6, activant la voie NF- κ B, et ainsi la production de cytokines pro-inflammatoires telles que le facteur de nécrose tumorale (Tumor Necrosis Factor, TNF- α), et les interleukines 1 (IL-1), 6 (IL-6) et 10 (IL-10). Le parasite va également stimuler la protéine Rab-5, induisant l'internalisation de *T. cruzi* par les macrophages (83). Les GPIL, quant à eux, vont activer le TLR4. Ce dernier, à travers la réponse primaire de différenciation myéloïde 88 (MyD88), va également activer la voie NF- κ B, et promouvoir la production des cytokines précédemment citées, ainsi que d'IL-12 et des interférons de type 1. La même réaction est induite par l'activation des TLR 7 et 9, activés par l'ADN du parasite, ce dernier contenant une grande quantité d'oligodeoxynucléotide CpG (courtes séquences d'ARN simple brin riches en cytosines et guanines méthylés), ainsi que par ses fragments d'ARN (82,84) (figure 17). Dans des modèles expérimentaux, l'inactivation génique (knock-out) du TLR2, TLR4, TLR9 et de la MyD88 ont amplifié l'infection par *T. cruzi*, prouvant leur importance dans le contrôle de la parasitémie (85–87).

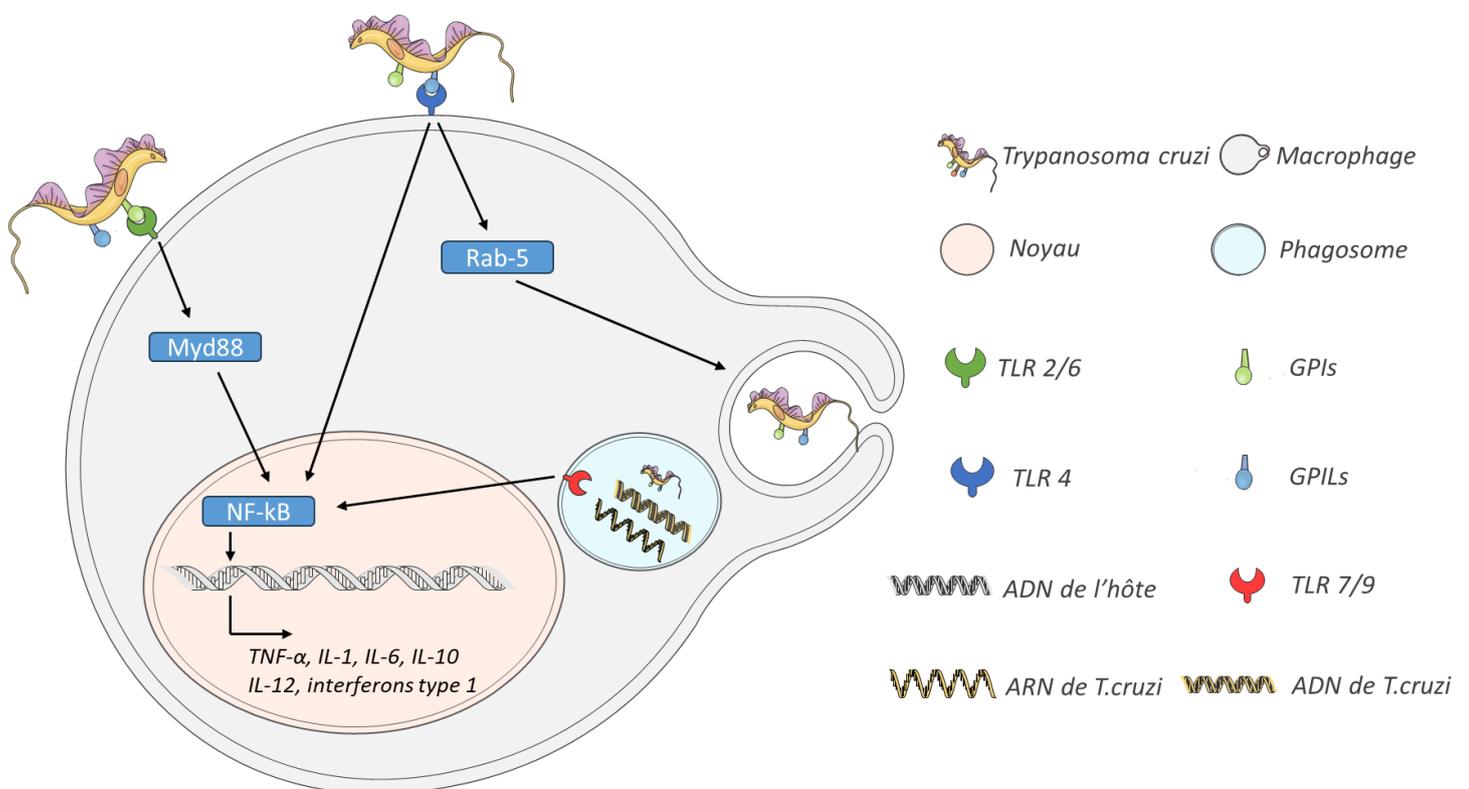


Figure 17. Activation de la réponse immunitaire innée via les TLR dans un macrophage.

Des mécanismes d'activation de la réponse immunitaire après l'infection par *T. cruzi* indépendants des TLR ont été observés, mais sont à l'heure actuelle méconnus. Cette activation peut se faire via les récepteurs NOD (NLR, NOD-like receptors), des récepteurs cytoplasmiques. D'autres études ont montré l'implication du facteur de transcription NFATc1, un membre de la famille des NFAT (Nuclear Factor of Activated T-cell), stimulé par la voie du Ca^{2+} . Ces mécanismes conduisent eux également à la production de cytokines pro-inflammatoires (88) (figure 18).

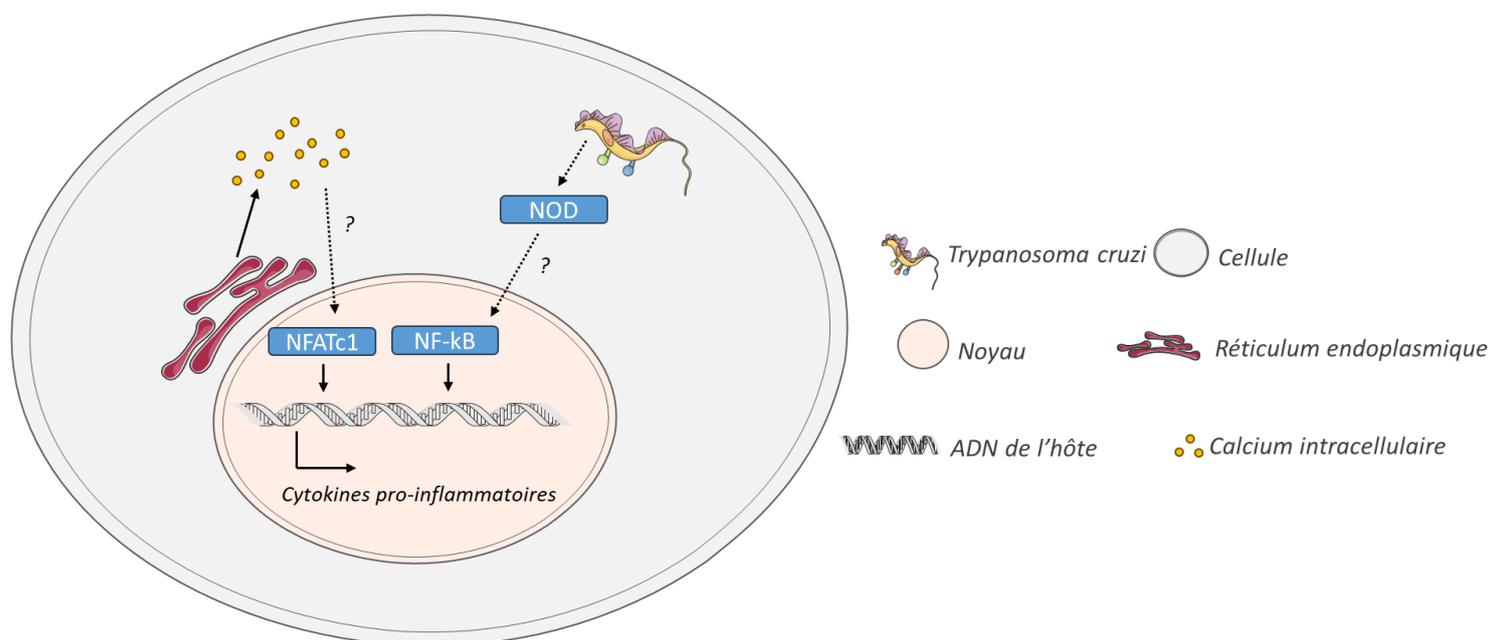


Figure 18. Activation de la réponse immunitaire innée indépendamment des TLR.

3.1.3. La destruction du parasite

Les différentes cytokines produites suite à l'activation de la voie NF-κB vont jouer un rôle dans la défense immunitaire, en promouvant l'inflammation et le recrutement d'autres cellules immunitaires au niveau du site d'infection. L'IL-12 va stimuler les cellules Natural Killers (NK), qui vont produire de l'interféron de type 2 (IFN- γ , interféron gamma), ainsi que du TNF- α . Ces deux cytokines vont alors activer les cellules phagocytaires, induisant la production d'espèces réactives à l'oxygène (ROS, Reactive Oxygen Species) et à l'azote (RNS, Reactive Nitrogen Species). Ce phénomène se nomme l'explosion respiratoire ou oxydative (oxidative burst). Les ROS sont des espèces chimiques oxygénées comprenant, entre autres, l'ion superoxyde ($O_2^{\cdot-}$), le peroxyde d'hydrogène (H_2O_2) et le radical hydroxyle HO. Les ROS sont produits par différentes sources endogènes, incluant la chaîne de transport d'électrons mitochondriaux et la NADPH (transporteur d'électron) (89).

Les RNS sont des espèces chimiques dérivant de l'oxyde nitrique ($^{\circ}\text{NO}$) et de l'ion superoxyde. Ils sont produits par la NADPH oxydase et l'oxyde nitrique synthase 2 (NOS2, Nitric Oxide Synthase).

Directement après l'infection, la phagocytose de *T. cruzi* par les macrophages (non activés) stimule la NADPH associée à la membrane, générant de grandes quantités d' $\text{O}_2^{\cdot-}$ dans le phagolysosome. L'activation des TLR par le parasite conduit à un rapprochement entre les mitochondries et le phagolysosome, permettant aux ROS mitochondriaux de traverser la paroi du phagolysosome et de participer à l'explosion oxydative (90). Toutefois, la concentration d' $\text{O}_2^{\cdot-}$ est souvent insuffisante pour éliminer le parasite. Lorsque les cytokines pro-inflammatoires ($\text{IFN-}\gamma$ et $\text{TNF-}\alpha$) activent les cellules phagocytaires, le facteur de transcription NF- κB induit la production de l'enzyme iNOS (codée par le gène NOS2). Cette dernière produit alors de grandes quantités de $^{\circ}\text{NO}$. Il réagit avec l' $\text{O}_2^{\cdot-}$ présent dans le phagolysosome pour produire du peroxy-nitrite (ONOO^-). ONOO^- et l' $\text{O}_2^{\cdot-}$ oxydent à la fois les protéines et l'ADN, altérant les liaisons chimiques qui maintiennent leur structure, et détruisent ainsi le parasite (91). L'apport des RNS dans la destruction de *T. cruzi* est important, des souris n'exprimant pas le gène NOS2 ayant une parasitémie significativement plus importante que les contrôles non mutés (92).

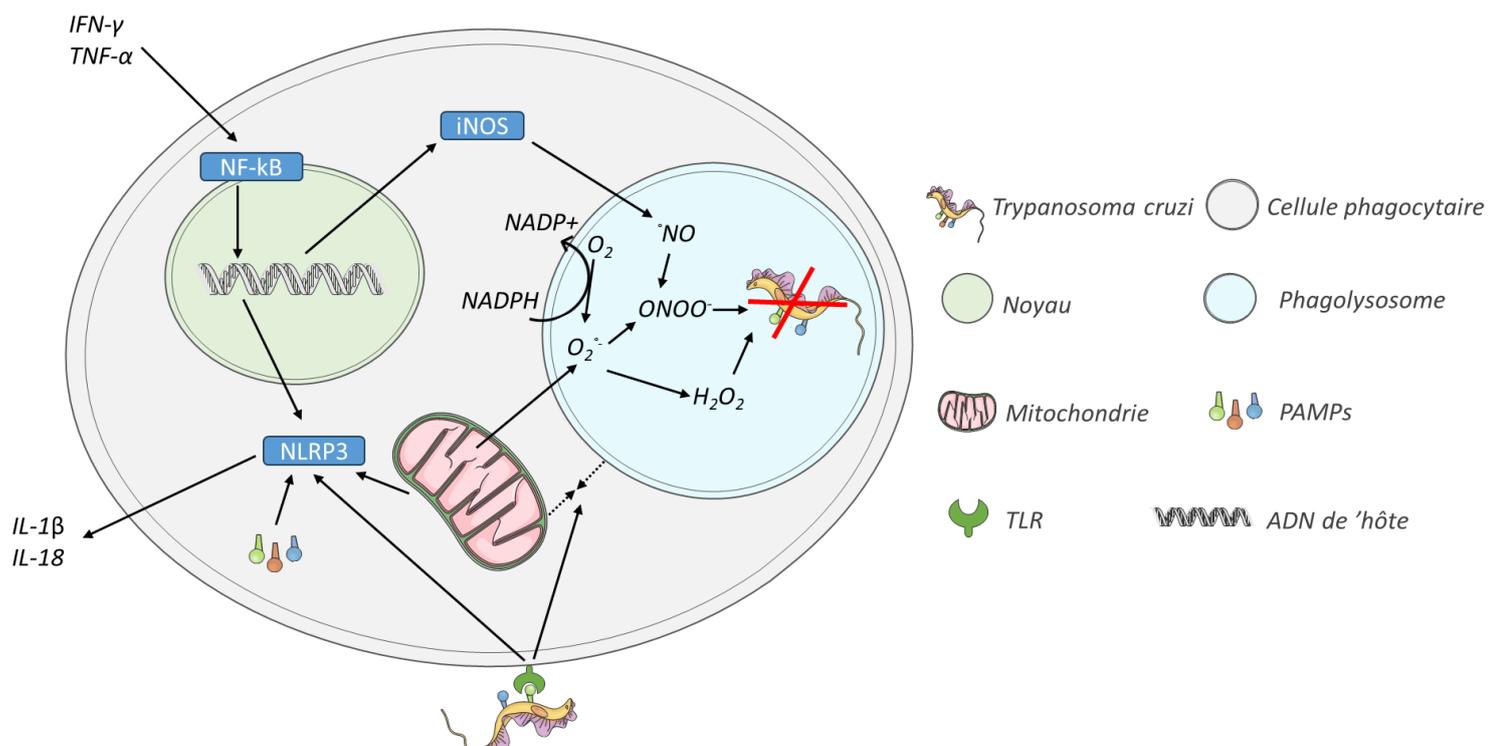


Figure 19. Mise en place des méthodes de destruction du parasite par les cellules phagocytaires.

Par ailleurs, la production de ROS mitochondriaux, l'activation des TLR et la stimulation des récepteurs NLRs participent à la mise en place de l'inflammasome NRLP3. Il s'agit d'un complexe protéique conduisant à la production d'interleukine IL-1 β , permettant le recrutement d'autres cellules immunitaires, ainsi que d'IL-18, favorisant une réponse des lymphocytes Th1. De plus, l'inflammasome est impliqué dans la pyroptose, une mort cellulaire rapide des cellules phagocytaires infectées, détruisant les parasites s'étant échappés du phagolysosome (90) ([figure 19](#)).

3.2. La réponse immunitaire adaptative

La réponse immunitaire adaptative se met en place plusieurs heures après la réponse immunitaire innée. Elle est spécifique à des antigènes précis.

3.2.1. Le déclenchement de la réponse immunitaire adaptative

La réponse immunitaire adaptative débute grâce à deux approches distinctes et complémentaires, qui engagent différents types cellulaires. La première voie implique l'activation des lymphocytes T CD4⁺ par l'intermédiaire de cellules présentatrices d'antigènes (APC), tandis que la seconde voie repose sur l'activation directe des lymphocytes B (LB) par les antigènes du parasite.

- L'activation des lymphocytes T par les APC

Les APC sont des cellules immunitaires incluant les macrophages et les cellules dendritiques qui exposent des fragments peptidiques du pathogène aux autres cellules immunitaires via le complexe d'histocompatibilité (MHC, Major Histocompatibility Complex), également appelé HLA (Human Leucocyte Antigen) chez l'être humain. Le MHC est un groupe de protéines présentes à la surface des cellules. Il joue un rôle crucial dans la reconnaissance des cellules propres et étrangères par le système immunitaire. Il existe deux types de MHC : le MHC de classe I (exprimé sur la plupart des cellules) et le MHC de classe II (exprimé principalement sur les APC). Lors de la phagocytose, des peptides du pathogène, ici *T. cruzi*, sont combinées aux protéines des MHC des APC.

Ces dernières, et principalement les cellules dendritiques, vont se déplacer vers le ganglion lymphatique (lymph node) le plus proche. Durant ce déplacement, elles deviennent matures et expriment en surface le complexe MHC/antigène (93)).

Les lymphocytes T (LT) sont des leucocytes produits dans la moelle osseuse et mûrissent dans le thymus. Il existe plusieurs types de LT, incluant les lymphocytes T auxiliaires (CD4⁺ ou T helper), les lymphocytes T cytotoxiques (CD8⁺ ou T killer) et les lymphocytes T régulateurs (Treg). Après maturation, les LT se dirigent dans les ganglions lymphatiques dans l'attente d'être activés. Lorsque les APC arrivent dans les ganglions lymphatiques, elles stimulent les lymphocytes T CD4⁺ via leur récepteur TCR (T-Cell Receptor) (94). La spécificité du TCR pour l'antigène est située dans le segment V, divisé en trois régions hypervariables : CDR1, CDR2 et CDR3. Elle est acquise lors du développement des lymphocytes dans le thymus : les segments V (variable), D (diversity) et J (joining) du gène TCR β et V et J du gène TCR α effectuent une recombinaison aléatoire, créant une multitude de clones différents, pouvant fixer des millions d'antigènes potentiels (95).

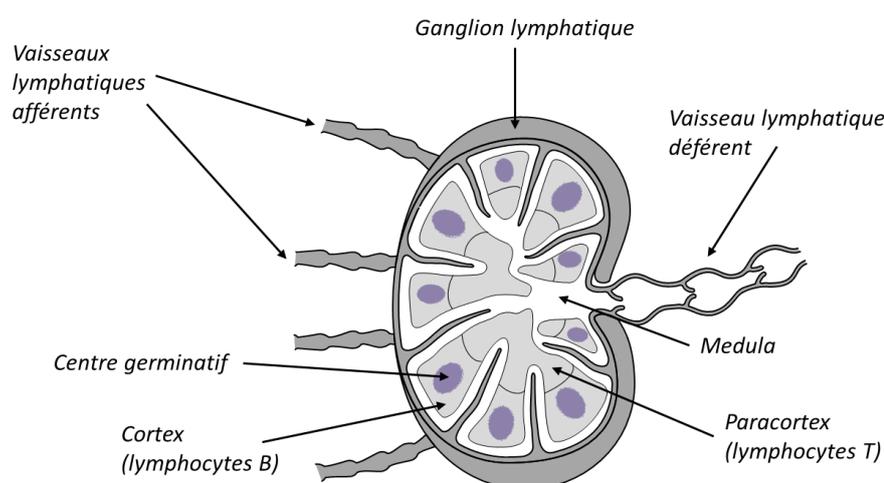


Figure 20. Représentation d'un ganglion lymphatique.

Les lymphocytes T interagissent avec les APC dans le paracortex des ganglions lymphatiques (figure 20). Suite à leurs interactions avec le complexe MHC-II/antigène, les lymphocytes T CD4⁺ immatures se spécialisent en différents types cellulaires, dont les Th1, Th2, Th17 et Treg. Cette différenciation est influencée par des signaux de régulation tels que l'IL-12 et l'IL-18. Dans le cas présent, ces cytokines sont favorisées par l'activation des TLR (récepteurs Toll-like) et par l'inflammasome NLRP3, qui activent la voie NF- κ B. Cette cascade de signaux tend à encourager la différenciation en lymphocytes Th1 par rapport aux lymphocytes Th2.

Bien qu'il soit crucial d'obtenir un équilibre entre les profils lymphocytes Th1 et Th2 pour permettre une inflammation efficace tout en maintenant un contrôle approprié, une prédominance de lymphocytes Th1 demeure préférable pour une réponse robuste contre *T. cruzi* (96). Les lymphocytes T CD8⁺, quant à eux, reconnaissent et sont activés par le complexe MHC-I/antigène des APC. Leur activation est également favorisée par la présence d'interféron gamma et d'IL-12, fortement exprimés lors de la réponse immunitaire innée à *T. cruzi* (95). Une fois activés, ces différents LT vont se répliquer en masse, puis se disséminer à travers l'organisme.

- L'activation des lymphocytes B par les antigènes de *T. cruzi*

Les lymphocytes B (LB) sont des leucocytes produits et développés dans la moelle osseuse. Les lymphocytes B immatures vont également se diriger vers les cortex des ganglions lymphatiques. Ils peuvent être activés de manière indépendante ou dépendante des lymphocytes T helper. Dans le premier cas, ils vont être directement activés par les antigènes circulants dans la lymphe. Ils sont capables d'ingérer les éléments pathogènes afin de les présenter sur leur MHC-II, et sont donc des APC. Dans le deuxième cas, après le premier contact avec les peptides pathogènes, ils interagissent avec les lymphocytes Th matures qui maximisent leur activation.

Une fois activés, les LB immatures effectuent des hypermutations somatiques dans les centres germinatifs. Il s'agit de mutations aléatoires dans la structure des BCR, composés d'une chaîne lourde et d'une chaîne légère. Les séquences ayant la meilleure affinité avec l'antigène présenté sont conservées, conduisant à une meilleure affinité des récepteurs avec les antigènes présentés, et donc une réponse spécifique et efficace. Après sélection, les LB se différencient en LB mémoires et migrent dans le cortex, ou en plasmocytes, producteurs d'immunoglobulines (Ig). Pour ces derniers, une dernière étape de maturation a lieu : la commutation isotypique (class switching), soit le changement de classes des Ig produites (**figure 21**). Les Ig, ou anticorps, sont des molécules produites par les LB qui vont se fixer à un pathogène ou à une cellule infectée pour favoriser son élimination par les cellules cytotoxiques. Il existe 5 classes majeures d'anticorps, appelées IgG, IgM, IgA, IgD et IgE. Ces 5 types d'Ig ont chacun une fonction bien précise (97).

Les IgG, les anticorps majoritairement présents dans le sérum (80%), sont les plus produits lors de l'infection par *T. cruzi* (98,99). Après la commutation isotypique, les plasmocytes se dirigent dans le cortex, où ils vont se répliquer et se disséminer dans l'organisme.

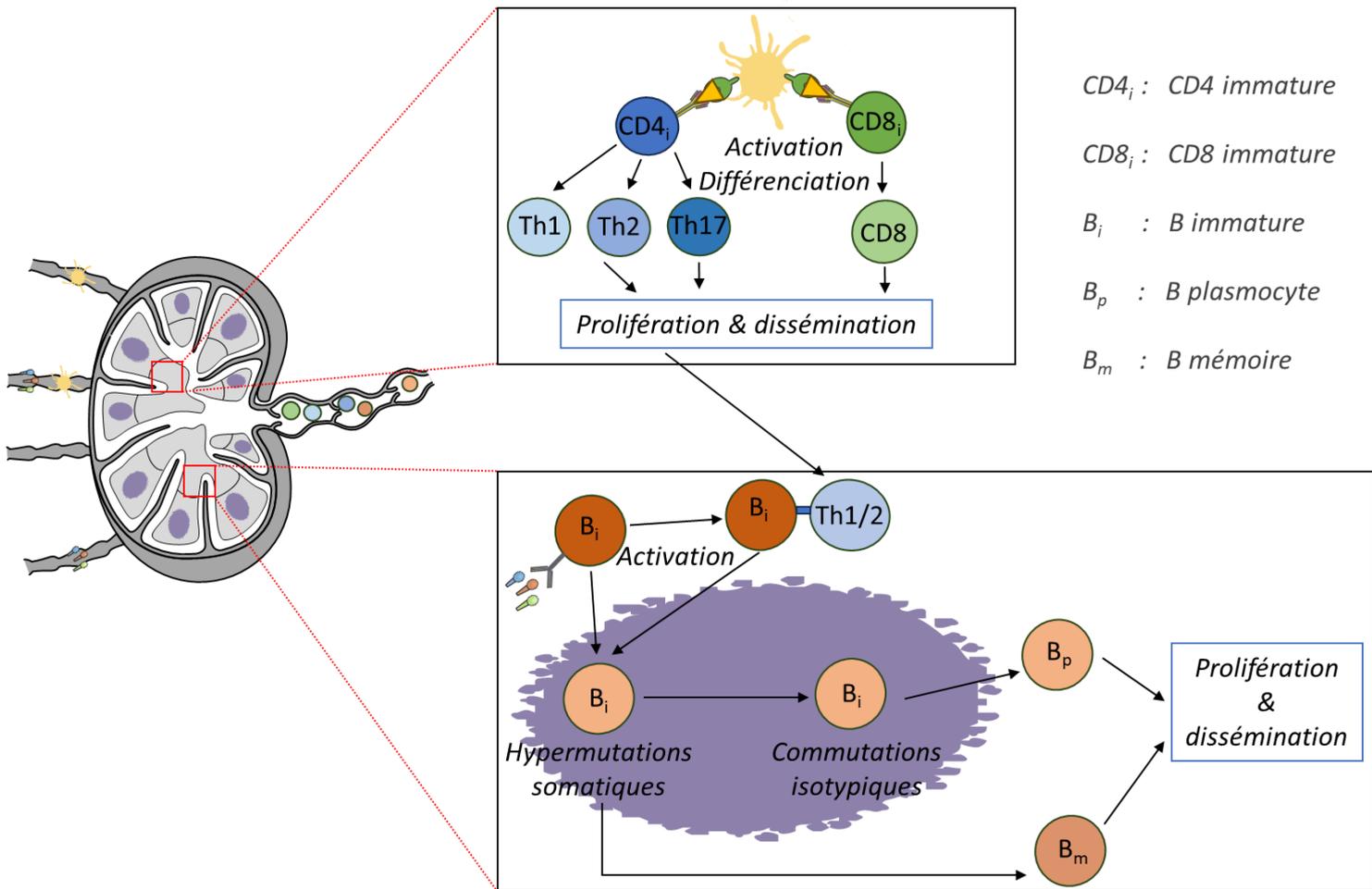


Figure 21. Représentation simplifiée de l'activation des lymphocytes B et T dans les ganglions lymphatiques.

3.2.2. La réponse immunitaire adaptative contre *T. cruzi*

Les différents lymphocytes B et T vont avoir des actions distinctes et complémentaires dans l'élimination du parasite. En effet, les plasmocytes produisent des anticorps ciblant le pathogène ou les cellules infectées. La fixation des anticorps sur les antigènes de *T. cruzi* favorise son élimination par les cellules phagocytaires. Concernant les cellules infectées, elles sont ciblées par les lymphocytes T CD8⁺ cytotoxiques qui les détruisent. Les lymphocytes T CD4⁺, quant à eux, jouent un rôle central de médiateur de la réponse immunitaire, produisant des cytokines induisant les réponses lymphocytes T CD8⁺ et plasmocytes.

- Le rôle des lymphocytes T CD4⁺

Les lymphocytes T CD4⁺ sont activés par divers types de cellules présentatrices d'antigènes, qu'elles soient des macrophages, des cellules dendritiques ou des lymphocytes B. L'importance de cette stimulation a été mise en évidence dans des expériences sur des modèles animaux. En effet, chez les souris infectées qui présentent une déficience en lymphocytes CD4⁺, on observe une parasitémie plus élevée par rapport aux souris infectées témoins, et toutes les souris atteintes finissent par succomber à l'infection (100). L'isolation de lymphocytes T CD4⁺ provenant de souris infectées et stimulées par des antigènes de *T. cruzi* a mis en évidence des profils distincts : des lymphocytes Th1 producteurs d'IFN- γ , et des lymphocytes Th2 producteurs d'IL-4 et d'IL-10. Toutefois, seuls les lymphocytes Th1 semblent avoir un impact contre le parasite in vitro, la production d'IFN- γ activant les macrophages et stimulant la génération de RNS détruisant le parasite (101) (figure 22). En revanche, l'importance des lymphocytes Th2 est non négligeable, régulant l'inflammation et ainsi les dommages tissulaires en résultant. Dans des modèles expérimentaux, l'inhibition d'IL-10, principalement produite par les lymphocytes Th2, conduit à une production plus élevée d'IFN- γ , de TNF- α et d'IL-12, à une plus faible parasitémie, mais également à une mort plus précoce (102)). Il est donc primordial de conserver une balance équilibrée entre les cytokines pro-inflammatoires (IFN- γ , TNF- α) et anti-inflammatoires (IL-10) (103). Les lymphocytes Th17 ont également un rôle dans la promotion de la réaction inflammatoire. L'inhibition d'IL-17A, dont ils sont les principaux producteurs, conduit à une diminution des cytokines pro-inflammatoires, et à une augmentation de la parasitémie (104). Les lymphocytes Th17 pourraient être plus efficaces que les lymphocytes Th1 dans l'élimination de *T. cruzi*, stimulant davantage la production de ROS par les macrophages et les lymphocytes T CD8⁺ in vitro (105).

- Le rôle des lymphocytes T CD8⁺

Les lymphocytes T CD8⁺ contribuent à l'élimination du parasite, leur inhibition dans des modèles expérimentaux conduisant à une parasitémie plus élevée (100). Ils agissent d'une part de manière indirecte, produisant d'une part de l'IFN- γ et du TNF- α , ce qui renforce l'activation des macrophages et la production de ROS/RNS et d'autre part, de l'IL-2, favorisant la prolifération des autres lymphocytes T.

De plus, ils éliminent directement le parasite ou les cellules infectées en produisant des granules cytolytiques, de petites vésicules contenant des protéines (**figure 22**). Parmi elles, la perforine va créer des pores dans la membrane de la cellule infectée, libérant des granzymes qui induisent la mort cellulaire (106). L'importance de la destruction des cellules infectées dans le contrôle de la parasitémie a été démontrée dans des modèles animaux. L'inhibition de la protéine $\beta 2m$ ($\beta 2$ microglobuline), qui est associée à la formation du complexe MHC-I/antigène, a conduit à une parasitémie plus élevée chez les souris infectées et à une mortalité accrue (107).

- Le rôle des lymphocytes B

L'importance des LB dans la résistance contre *T. cruzi* a déjà été prouvée (98,108), leur inhibition induisant une réduction de la concentration des lymphocytes T CD4⁺ et T CD8⁺ (109), et notamment des lymphocytes T Th1 (110) (**figure 22**). Peu après l'infection, une très forte réponse polyclonale apparaît, produisant une grande variété d'anticorps non spécifiques (anergie), principalement des IgG (98,99), induisant une hypergammaglobulinémie (augmentation anormalement élevée des IgG). Cela résulte en une réponse spécifique retardée. Une étude portant sur deux lignées de souris, une résistante et l'autre sensible, a montré que les souris résistantes ont une réponse B délayée, avec une expression d'IgG plus faible et mixte, mais à terme une réponse plus spécifique. Elles présentent également l'expression d'IFN- γ et d'IL-6 plus précoce. À l'inverse, les lignées sensibles présentent une hypergammaglobulinémie avec une production d'IgG1 et d'IgG2a, et une réponse tardive orientée Th2 (111). Par ailleurs, la production de lymphocytes B semble altérée par *T. cruzi*, développant un mécanisme d'immunosuppression induisant l'apoptose précoce des cellules B (112,113) et leur diminution dans la moelle (114). Ainsi, la non-spécificité des anticorps pourrait être due à la mort précoce des LB, ou à la grande variété des antigènes de *T. cruzi* (115,116), notamment les GIPL (117). Dans les deux cas, ces mécanismes pourraient être mis en place directement par *T. cruzi*, la manipulation de la réponse B par des pathogènes, qu'il s'agisse de parasites, de virus ou de bactéries, ayant déjà été observée dans d'autres pathologies (118).

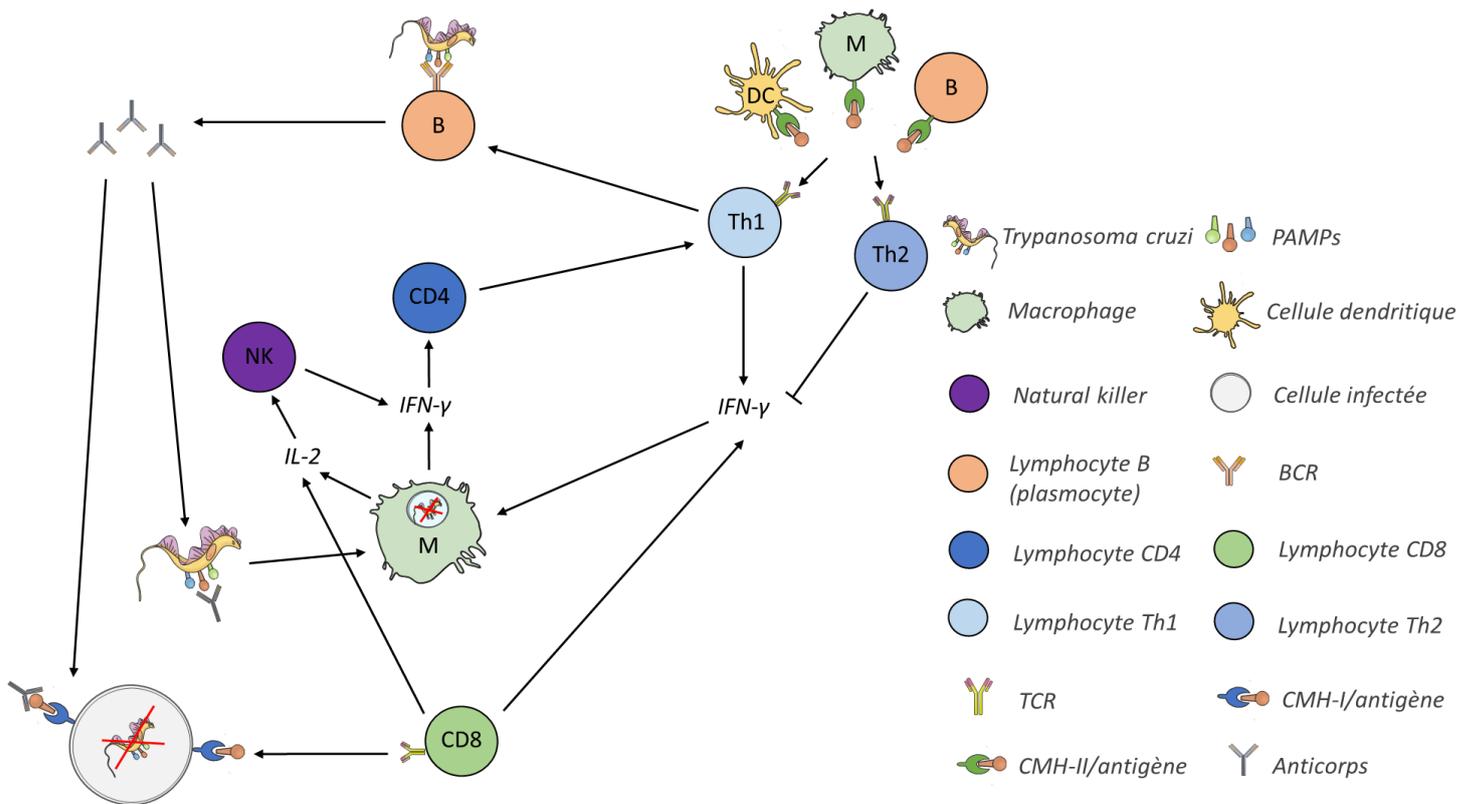


Figure 22. Représentation simplifiée de la réponse immunitaire adaptative contre *T. cruzi*.

3.3. L'évasion de *T. cruzi* face à la réaction immunitaire

Bien que différents mécanismes de défense soient mis en place par l'hôte infecté, ils ne sont pas toujours suffisants pour venir totalement à bout de l'infection. En effet, *T. cruzi* peut parvenir à s'échapper du phagolysosome des cellules phagocytaires afin d'atteindre le cytoplasme et se répliquer. De plus, il produit des enzymes antioxydantes, associées à la virulence du parasite (119), réduisant l'efficacité du peroxy-nitrite (ONOO^-). Le parasite semble notamment agir sur le fonctionnement de la chaîne mitochondriale, changeant l'activité des complexes afin de favoriser la production de H_2O_2 (120), déséquilibrant la balance ROS/RNS et donc la production de peroxy-nitrite (91).

En parallèle, plusieurs études suggèrent que les ROS favorisent l'infection de *T. cruzi* dans les macrophages (121). Il serait notamment capable de promouvoir la production de ROS en réprimant le facteur de transcription NFE2L2 (NRF2), régulateur de gènes ayant une activité antioxydante (122). NFE2L2 a été observé sous-exprimé dans les CCC par rapport à des contrôles sains (123).

Après avoir réussi à échapper à la réponse immunitaire innée, le parasite peut aller infecter d'autres organes, tels que le cœur et le système digestif. L'infection de cardiomyocytes de souris par *T. cruzi* altère le potentiel de membrane des mitochondries et par conséquent le transport d'électrons, augmentant la production de ROS, d'IL1 β , de TNF- α et d'IFN- γ (124), ces cytokines étant associées au développement des CCC (121).

Par ailleurs, *T. cruzi* pourrait influencer sur la réponse immunitaire adaptative. D'une part, comme détaillé précédemment, il semble promouvoir une forte production de lymphocytes B non spécifiques, ce qui réduit leur efficacité. D'autre part, il serait en mesure de limiter l'action des lymphocytes T. Des études in vitro ont montré que la cruzipain, une protéase sécrétée par le parasite, perturbe la voie de signalisation NF-kB dans les premières phases de l'infection (moins d'une heure), entraînant une diminution de la production d'IL-12 (125), importante pour la différenciation des lymphocytes T CD4⁺ en Th1. De même, l'expression de différentes enzymes de *T. cruzi*, telles que les trans-sialidases et la mucine, contribue à réduire la sécrétion d'IL-2 et d'IFN- γ chez la souris, entraînant une inhibition de la production de lymphocytes T. Notamment, une augmentation de la concentration d'IL-10 induite par les trans-sialidases favorise la production de lymphocytes Th2 aux dépens des Th1, et ainsi à une diminution de la réaction inflammatoire (126,127). En outre, il a été expérimentalement constaté que la stimulation de lymphocytes B par *T. cruzi* conduit à la production d'IL-10. Cela entraîne une diminution des lymphocytes T CD4⁺ et une augmentation de leur apoptose (128).

Il est important de noter que la modulation de la réponse immunitaire peut diverger selon les souches du parasite. Une illustration de ce phénomène se trouve dans une étude comparative entre deux souches appartenant aux groupes DTU-I et DTU-II, où une réponse inflammatoire plus prononcée a été observée in vitro chez les souches DTU-II, tandis que les souches DTU-I ont démontré une augmentation plus marquée de la réponse anti-inflammatoire stimulée par les IL-10 (129).

Ainsi, *T. cruzi* tend à moduler la réponse immunitaire de l'hôte par différents mécanismes afin d'assurer sa survie, et même de promouvoir sa réplication. Ces différentes actions conduisent à une persistance du parasite, et favorisent sa dissémination dans l'organisme de l'hôte.

3.4. Le développement des formes cardiaques

Le lien entre la persistance du parasite et le développement des formes chroniques, notamment des CCC, est controversé. En effet, des études menées sur la souris ont conduit à des conclusions opposées, la charge parasitaire étant parfois augmentée, parfois réduite lors du développement de symptômes cardiaques (121). Cependant, la forte réponse inflammatoire observée en phase aiguë est maintenue dans le cœur des hôtes atteints de cardiomyopathies chagasiques, entraînant des dommages irréversibles aux tissus.

3.4.1. Une absence de régulation de la réaction immunitaire

- Un déficit dans la réponse anti-inflammatoire

Le processus pathogénique conduisant au développement des CCC demeure largement méconnu. Toutefois, il est évident qu'une réponse inflammatoire exacerbée persiste dans les tissus cardiaques, caractérisée par un infiltrat de macrophages, lymphocytes et plasmocytes (45). De nombreuses études, comparant des sujets atteints de CCC à des asymptomatiques (ASY), ont associé cette forte inflammation à un déséquilibre entre les réponses immunitaires inflammatoires et anti-inflammatoires. En effet, des cytokines pro-inflammatoires spécifiques des lymphocytes Th1, telles que l'IFN- γ et du TNF- α , sont exprimées chez tous les patients en phase chronique. Cependant, tandis que les patients ASY maintiennent un équilibre dans le profil des lymphocytes Th1/Th2, les patients atteints de CCC montrent une accumulation de lymphocytes Th1. En parallèle, la concentration en lymphocytes T régulateurs (Treg) est plus faible chez les patients CCC (96,130).

L'interleukine 10 (IL-10) est un régulateur clé de l'IFN- γ . Elle est majoritairement produite par certains Treg, mais également par les lymphocytes B. Bien que la concentration en lymphocytes B soit similaire chez les ASY et les CCC, leur stimulation par des protéines de *T. cruzi* induit la production d'IL-10 et de TNF- α chez les ASY, tandis que seul du TNF- α est produit chez les CCC (131). Différentes études ont confirmé que la concentration en IFN- γ et TNF- α est plus forte chez les patients atteints de CCC (132), et la concentration d'IL-10 plus faible (130) (**figure 23**).

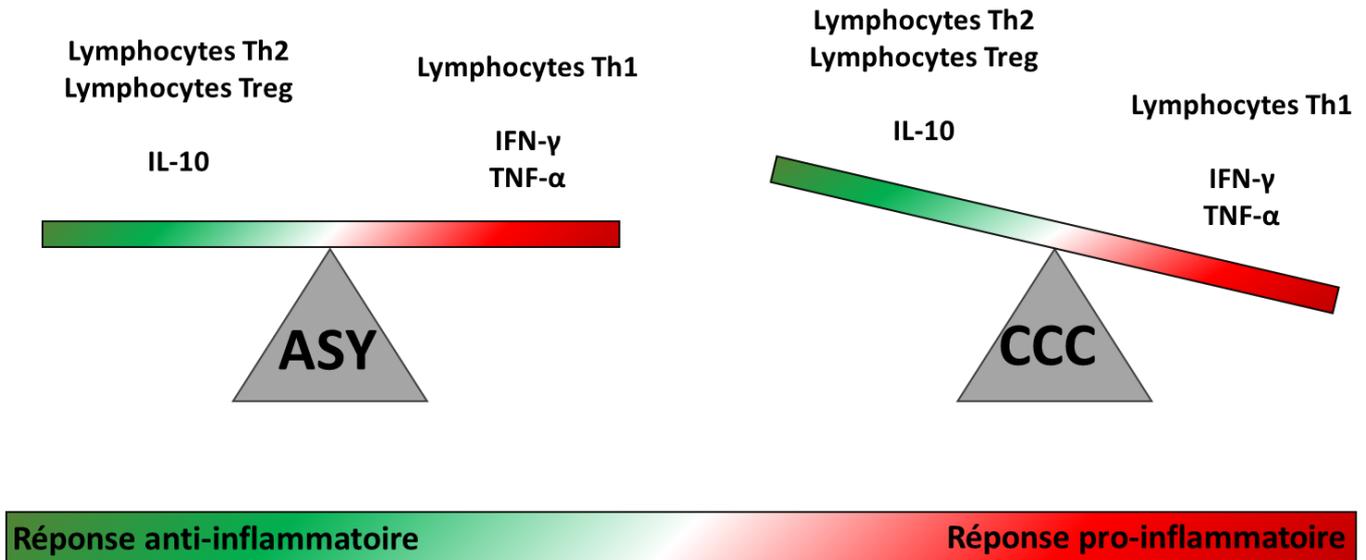


Figure 23. Déséquilibre dans la réponse pro- et anti-inflammatoire chez les patients atteints de CCC par rapport aux patients ASY.

- De la réaction immunitaire aux dommages cardiaques

Les lymphocytes T CD4⁺ peuvent également se différencier en lymphocytes T Th17, potentiellement liés aux dommages cardiaques. En effet, la concentration d'IL-17, principalement produite par les lymphocytes Th17, est plus élevée chez les ASY que chez les CCC. Une faible expression d'IL-17 est corrélée avec la détérioration de la fonction cardiaque, notamment de la fraction d'éjection cardiaque du ventricule gauche (133). Chez la souris, l'inhibition d'IL-17 conduit à une réduction de la charge parasitaire dans le tissu cardiaque, mais également à plus forte myocardite et à une mort prématurée (134).

Dans le tissu cardiaque des CCC, une forte concentration de lymphocytes T CD8⁺ et Th1 a été observée. Ces deux types cellulaires produisent de l'IFN-γ, qui active les macrophages et stimule la production de °NO. L'augmentation de la concentration de °NO a été associée à une aggravation des CCC, de même que la diminution de glutathionne peroxydase et de superoxyde dismutase (SOD), des anti-oxydants (135). En parallèle, des dysfonctionnements mitochondriaux ont été observés dans le sérum de patients atteints de CCC et dans le tissu cardiaque souris infectées. Cela se traduit par une diminution de l'activité de la chaîne respiratoire, notamment des complexes I et III, principaux producteurs d'O2[°]. Le stress oxydatif induit par l'accumulation d'O2[°] provoque des dommages à la fois au niveau des mitochondries et du tissu cardiaque (136,137).

De plus, la diminution de la production d'ATP causée par les dysfonctionnements mitochondriaux contribue à la détérioration de la fonction cardiaque, les cardiomyocytes n'ayant pas assez d'énergie pour maintenir une contraction adéquate.

Par ailleurs, les lymphocytes T CD8⁺ pourraient directement contribuer aux dommages tissulaires. L'inhibition du MHC-I de souris infectées montre une absence de T CD8⁺ et d'inflammation dans les tissus cardiaques, suggérant leur rôle dans la destruction des cellules de l'hôte (107). De plus, une différence de profil des T CD8⁺ apparaît durant l'infection par le parasite. Alors que les T CD8⁺ inflammatoires (producteurs d'IFN- γ) sont majoritaires durant les 60 premiers jours post-infection, la quantité de T CD8⁺ cytotoxiques (producteurs de perforine) présents dans le tissu cardiaque augmente drastiquement dans les stades plus tardifs. Les lésions cardiaques ont par ailleurs été associées aux T CD8⁺ cytotoxiques, et non aux T CD8⁺ inflammatoires (138). Des résultats similaires ont été obtenus chez des patients atteints de la maladie de Chagas. En effet, la stimulation de lymphocytes T CD8⁺ par des peptides de *T. cruzi* n'a entraîné aucune production d'IFN- γ chez des patients atteints de CCC sévère, alors que l'inverse a été observé chez des patients atteints de CCC modérée ou chez des individus sains (139). La concentration de lymphocytes T CD8⁺ producteurs de perforine et de granzyme B est par ailleurs plus importante chez les CCC que chez des patients ASY (140).

En résumé, diverses études ont mis en évidence des variations notables dans le profil immunitaire des cellules chez les patients en phase chronique. L'absence de régulation de la réponse inflammatoire, provoquée par une différenciation déséquilibrée des lymphocytes CD4⁺ ou par une réponse spécifique des lymphocytes B et CD8⁺, conduit à une accumulation de ROS et de RNS. Cette forte concentration d'espèces réactives entraîne des dommages irréversibles aux tissus cardiaques. La spécialisation des lymphocytes T CD4⁺ et T CD8⁺ pourrait également contribuer à la détérioration du cœur (**figure 24**). Cette inflammation chronique pourrait donc être liée au développement des CCC, même si les mécanismes sous-jacents demeurent encore peu compris.

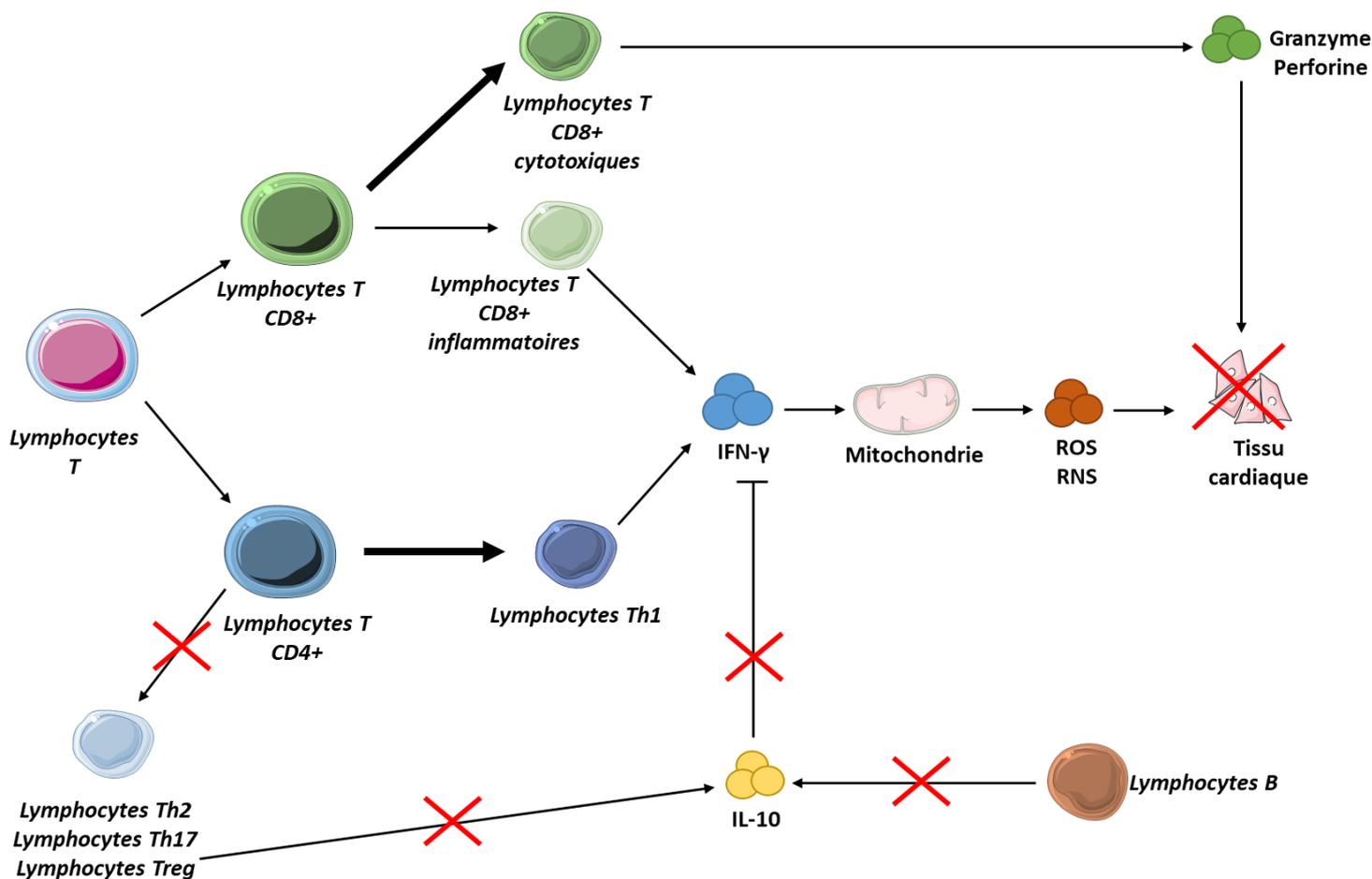


Figure 24. Schéma récapitulant les connaissances actuelles concernant le rôle de la réponse immunitaire dans le développement des lésions cardiaques observées dans les CCC.

3.4.2. Une réaction auto-immune

La réponse inflammatoire exacerbée dans le tissu cardiaque favorise le développement des CCC, mais les raisons de sa persistance demeurent incertaines. La première hypothèse suggère que la persistance du parasite provoque une stimulation continue des cellules immunitaires. Cependant, la variabilité des niveaux de parasites constatés aux différentes phases de la maladie, qu'elle soit asymptomatique ou symptomatique, remet en question cette supposition.

La deuxième possibilité est qu'une réaction auto-immune se développe, les cellules immunitaires détruisant directement les cellules de l'hôte, notamment les cardiomyocytes (141). De nombreuses études suggèrent que *T. cruzi* pourrait mettre en place un mimétisme des molécules de l'hôte. Par exemple, l'antigène B13 (AAAGDK), présent à la surface du parasite, est très proche

d'un épitope présent sur la chaîne lourde de la myosine cardiaque (AAAGDK). Cunha-Neto E. et al. ont par ailleurs démontré que tous les lymphocytes T obtenus à partir de patients atteints de CCC réagissent à la chaîne lourde de la myosine cardiaque et à l'antigène B13, contre 14% de celles des lymphocytes T issu de patients ASY (142). Une autre étude, portant sur la souris, a confirmé la réaction d'anticorps avec la myosine cardiaque. De la myosine purifiée a été extraite des muscles cardiaques de souris saines ou immunisées avec de la cruzipain, une protéine présente à la surface de *T. cruzi*. Des tests de précipitations ont alors montré que les anticorps des souris immunisées se fixaient à la fois sur la cruzipain et la myosine enrichie. De plus, les cellules de la rate des souris préalablement immunisées ont proliféré en réponse à la stimulation par la myosine enrichie, contrairement aux cellules des souris témoins. (143). Des glycosphingolipides cardiaques de souris, qui sont des lipides impliqués dans la construction des membranes cellulaires, présentent également des similarités architecturales avec certaines molécules de *T. cruzi*. Un test ELISA a révélé que les anticorps des patients atteints de la maladie de Chagas réagissent à la fois avec les formes épimastigotes du parasite et avec les glycosphingolipides présents dans le cœur des souris (144).

Les anticorps présents dans le sérum des patients atteints de la maladie de Chagas semblent également cibler l'activité nerveuse, en particulier celle du cœur, via les récepteurs des neurotransmetteurs. Les récepteurs adrénergiques réagissent à l'adrénaline et la noradrénaline et sont impliqués dans l'augmentation de la fréquence cardiaque. Une étude menée sur des patients en phase aiguë et chronique (ASY et CCC), a montré que deux anticorps différents se lient spécifiquement aux récepteurs adrénergiques β_1 du cœur et β_2 de la rate chez le rat. De manière générale, l'activité des IgG anti- β_1 (se fixant sur les récepteurs adrénergiques β_1) est faible durant la phase aiguë, et augmente avec le temps d'infection. De plus, elles inhibent l'action de norépinephrine sur la contractibilité des oreillettes (145). À l'inverse, les récepteurs cholinergiques interagissent avec l'acétylcholine, et diminuent la fréquence cardiaque. Des IgG obtenus à partir du sérum de patients atteints de la maladie de Chagas semblent agir comme des activateurs des récepteurs cholinergiques, abaissant la fréquence cardiaque chez le rat et ainsi l'efficacité de la fonction cardiaque. La concentration de ces anticorps est notamment plus élevée chez les patients présentant des altérations du système nerveux autonomes que chez les ASY (146).

Des réactions entre du sérum de patients atteints de la maladie de Chagas et des molécules impliquées dans d'autres processus biologiques, tels que le système nerveux et la réponse immunitaire, ont également été observées (147).

La forte réaction polyclonale des lymphocytes B et le manque de spécificité des anticorps observés en phase aiguë pourraient expliquer ce phénomène, la sélection d'anticorps réactifs ayant lieu plus tardivement, lorsque le parasite a infiltré les cellules cardiaques. Récemment, Ricci AD. et al. ont construit un atlas de l'ensemble des anticorps détectés des individus infectés par différentes souches de *T. cruzi*, appartenant aux DTU I et II (148). Ces données pourraient à l'avenir participer à une meilleure caractérisation des anticorps auto-réactifs, et à la compréhension du mécanisme auto-immunitaire apparaissant chez les CCC.

Ainsi, la réponse immunitaire provoquée par l'infection de *T. cruzi* se caractérise par une forte inflammation, entraînant la production de cellules Th1, de T CD8⁺ et de plasmocytes. Cette réponse, combinée à la production de ROS/RNS par les macrophages, permet la destruction du parasite. Cependant, chez les CCC, l'aggravation de la réaction inflammatoire semble causer des dommages cardiaques irréversibles. Les mécanismes responsables de la persistance de cette réponse demeurent encore largement inconnus, et de nombreuses études sont actuellement menées à différents niveaux moléculaires pour élucider ce phénomène.

Chapitre 4 : Etude de la maladie de Chagas par l'utilisation de différentes omiques

De nombreuses recherches ont abordé, bien que partiellement, la réponse immunitaire observée dans les formes aiguës et chroniques de la maladie de Chagas. Cependant, les facteurs sous-jacents à l'apparition de symptômes demeurent méconnus. Avec l'émergence de techniques d'analyse à grande échelle, il devient possible d'approfondir notre compréhension des mécanismes moléculaires et cellulaires à l'œuvre chez l'hôte. Ce chapitre explorera les diverses approches omiques utilisées jusqu'à présent pour comprendre le processus pathogénique propre aux CCC, ainsi que les conclusions qui en découlent. Il se concentrera sur les différents niveaux moléculaires impliqués dans l'expression et la régulation des éléments contribuant au développement des CCC.

4.1. Qu'est-ce que les "omiques" ?

Le terme "omique" est apparu dans les années 2000. Il fait référence à des domaines d'étude qui se concentrent sur l'analyse globale de divers composants biologiques. Il est dérivé du suffixe "-omique", qui désigne l'étude de molécules spécifiques (figure 25). Au cours des dernières années, l'avancée des technologies d'analyses à grande échelle a permis l'identification et la caractérisation d'une grande diversité de composants biologiques.

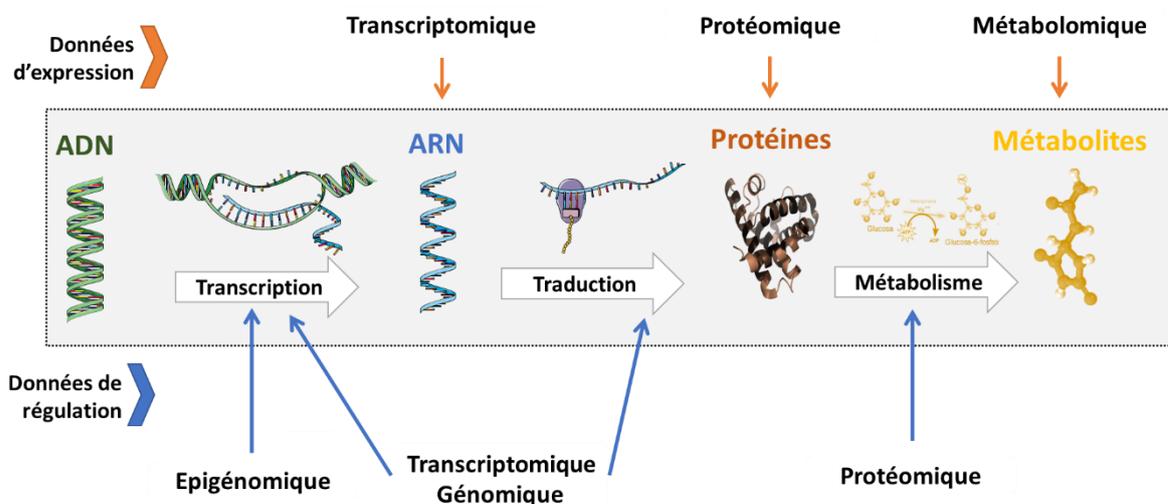


Figure 25. Présentation des principales "omiques" permettant d'analyser la globalité d'un processus biologique.

Les différentes approches omiques seront présentées dans cette section en suivant deux axes distincts. Le premier axe se penchera sur les omiques permettant d'étudier les molécules directement impliquées dans l'exécution des processus biologiques, englobant ainsi l'analyse des gènes codants, des protéines et des métabolites. Le deuxième axe, quant à lui, se concentrera sur les éléments génétiques ou épigénétiques qui ont pour principale fonction la régulation de l'expression et/ou de la fonction des gènes codants et des protéines.

4.1.1. Les molécules jouant un rôle actif dans les processus biologiques

Les informations relatives à l'expression génique, protéique et métabolique jouent un rôle essentiel dans la compréhension approfondie des mécanismes biologiques. La transcriptomique, qui consiste en l'étude de tous les éléments transcrits à partir de l'ADN, offre une vision globale de l'expression de l'ensemble des gènes au sein d'un organisme. Son importance dans la recherche sur les maladies a été largement démontrée. Par exemple, dans le contexte de l'autisme (149), elle a permis de mettre en évidence des altérations dans l'épissage alternatif du gène A2BP1 associé à cette pathologie, affectant potentiellement de nombreux gènes cibles. De même, elle a permis d'identifier des régulateurs clés associés au diabète de type 2 (150). Le niveau d'expression génique peut être mesuré via des puces à ADN (microarray), qui ciblent un ensemble de gènes prédéterminés (151). Le séquençage de l'ARN (RNA-sequencing), quant à lui, séquence directement les molécules d'ARN, permettant non seulement de quantifier l'expression de tous les gènes connus d'un organisme, mais également de détecter éventuellement de nouveaux transcrits, des isoformes, ou des variants (152). Récemment, des approches plus complexes ont été développées, mesurant l'expression génique à l'échelle cellulaire (single-cell) (153) ou spatiale (spatial-transcriptomique), ajoutant la position de cellules (ou de groupes de cellules) dans un tissu (154).

Les protéines sont des acteurs clé de la dynamique cellulaire et de la réponse aux signaux environnementaux. Du fait des différents mécanismes de régulation de la traduction, la transcription d'un gène n'est pas toujours corrélée à la production de sa protéine. À travers des méthodes de spectrométrie, notamment de spectrométrie de masse, la protéomique permet la quantification, la caractérisation et l'analyse fonctionnelle des protéines. Récemment, elle a permis d'identifier des protéines liées aux formes graves du Covid-19 associées à l'activation du complément, présentant des différences entre les individus ayant nécessité une hospitalisation, ou non (155).

Colak D. et al. ont mené une étude mêlant données transcriptomiques et protéomiques afin de comprendre le processus pathogénique associé aux cardiomyopathies dilatées (156). Ils ont identifié 1.211 gènes et 135 protéines associés à cette pathologie, dont seulement 16 en commun, démontrant le lien parfois flou entre expression génique et protéique.

Enfin, les métabolites sont des petites molécules organiques qui participent ou sont issues du métabolisme, tels que les glucides, les acides aminés ou les acides gras. Leur implication dans le développement de maladies n'est plus à démontrer, étant parfois directement responsables de certaines. Par exemple, dans le cas de la phénylcétonurie, un déficit en phénylalanine oxydase conduit à l'accumulation de phénylalanine, induisant des dommages cérébraux. Les méthodes appliquées en métabolomique varient considérablement en fonction des métabolites, étant donné leur diversité en termes de propriétés structurales. De nombreux métabolites, associés aux acides aminés et au cycle de l'urée, ont été associés au risque de maladies cardiovasculaires, et pourraient être utilisés comme biomarqueurs (157).

La quantification et caractérisation des gènes, protéines et métabolites permet d'avoir une vue d'ensemble des processus biologiques. Toutefois, leur étude seule ne suffit pas toujours à comprendre les perturbations biologiques observées dans des phénomènes d'intérêt.

4.1.2. Les éléments génétiques et épigénétiques impliqués dans les processus de régulation

Les données d'expression seules ne suffisent pas à expliquer l'apparition ou l'aggravation de processus pathogéniques. Bien que les gènes et les protéines jouent parfois un rôle régulateur, de nombreux autres éléments génétiques ou épigénétiques influent sur leur expression ou leur fonction. En effet, de nombreux éléments régulateurs viennent moduler leur production, et doivent être pris en considération dans l'étude de maladie. Parmi eux, la présence de mutations dans le génome peut moduler l'expression de gènes ou la fonctionnalité de protéines. À travers des méthodes de génotypage telles que la GWAS (Genome-Wide Association Study) ou le séquençage, comme le WES (Whole Exome Sequencing) ou le WGS (Whole Genome Sequencing), la génomique permet une étude approfondie de ces variations. Dans le cadre du cancer du sein, il existe une corrélation entre de multiples mutations dans les gènes BRCA1 et BRCA2 et le développement de cette maladie. Parallèlement, de nombreux autres SNP (Single Nucleotide Polymorphism) sont liés à ces gènes, amplifiant davantage l'association avec le cancer du sein (158).

Par ailleurs, des mutations dans le promoteur du gène de la globuline gamma semblent affecter la liaison à l'ADN des facteurs de transcription BCL11A et ZBTB7A, pouvant expliquer le développement d'hémoglobinopathies comme la drépanocytose (159).

En plus des altérations permanentes de l'ADN résultant des mutations, il existe d'autres modifications qui peuvent également influencer l'expression des gènes, et de fait le développement de maladies. L'épigénomique se focalise sur l'étude des altérations épigénétiques, soit des changements réversibles ne modifiant pas la séquence d'ADN. Par exemple, la méthylation de l'ADN ajoute un groupement méthyl sur une cytosine suivie d'une guanine, formant un îlot CpG. La méthylomique est souvent étudiée via des méthodes de bisulfite, convertissant les cytosines non méthylées en uracile. L'ADN est séquencé ou analysé via des puces (similaires aux puces à ADN), détectant la présence d'uraciles et donc la présence d'îlots CpG. La présence (ou absence) de méthylation dans des régions régulatrices, comme dans les promoteurs, peut avoir un impact dans la régulation de l'expression génétique, en modulant la fixation de facteurs de transcriptions (**figure 26**). Par exemple, la méthylation du promoteur du gène IRAK3 semble impliquée dans son inhibition dans les cancers du côlon, du poumon et de la prostate (160). L'impact de la méthylation dans les maladies cardiovasculaires a également été largement étudié, avec de nombreuses études démontrant des différences de méthylation dans des associés à l'activation de la transcription, les muscles squelettiques ou à la matrice extracellulaire, comme TEAD1 et PTPRN2 (161).

La structure de la chromatine peut également avoir un impact direct sur la régulation de l'expression génétique. En effet, son ouverture cela facilite l'accès à l'ADN des éléments régulateurs, tels que les facteurs de transcription aux gènes. Elle est enroulée autour d'un groupement de protéines appelées histones. L'acétylation de leur extrémité N-terminale permet la décompaction de la chromatine. L'ATAC-Seq (Assay for Transposase-Accessible Chromatin Using Sequencing) permet d'étudier directement sa structure, et le ChIP-Seq (Chromatin Immunoprecipitation followed by Sequencing) analyse les séquences d'ADN avec lesquelles des protéines interagissent. L'impact de la compaction de la chromatine dans le développement de maladies a déjà été prouvé, comme dans le cancer (162) ou les maladies cardiaques (163).

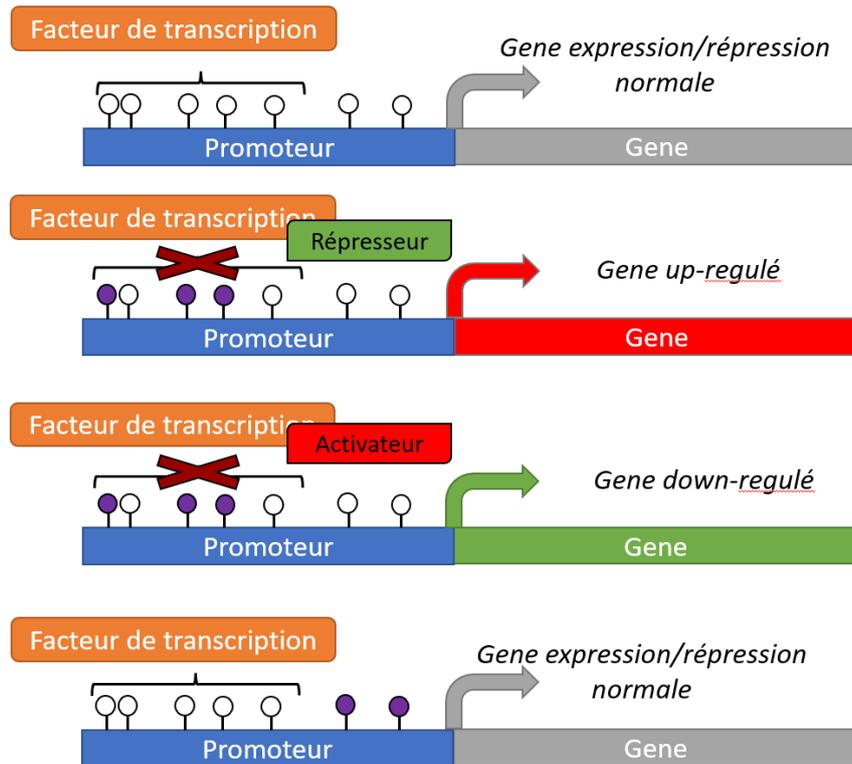


Figure 26. Impact de la méthylation de l'ADN (violet) dans la régulation de l'expression génique.

Enfin, la transcription d'ARN non codant, une catégorie d'ARN qui, contrairement aux ARN messagers, ne sont pas traduits en protéines, permet la régulation de l'expression génique et protéique. L'exploration de ces ARN se fait par le biais de techniques de transcriptomique, révélant ainsi leur niveau d'expression et leur influence potentielle dans la régulation génétique. Il existe une variété d'ARNs non codants, tels que les ARN ribosomiques (ARNr), les ARN de transferts (ARNt), les ARN circulaires (ARNcirc), les micro-ARN (miARN) et les ARN long non codants (ARNlnc). Ces deux derniers sont les plus connus, et les plus étudiés dans le cadre de pathologies. Les miARN sont des petits ARN simple brin d'une vingtaine de nucléotides qui régulent l'expression des gènes par complémentarité, partielle ou totale, avec la séquence d'ARN messager. Cette association se fait par le biais du complexe RISC (RNA-Induced Silencing Complex), qui conduit le miARN à sa cible (164) (figure 27). De nombreux miARN ont déjà été associés à des maladies, comme les maladies cardiaques, comme miR-133, associé à des infarctus du myocarde (165).

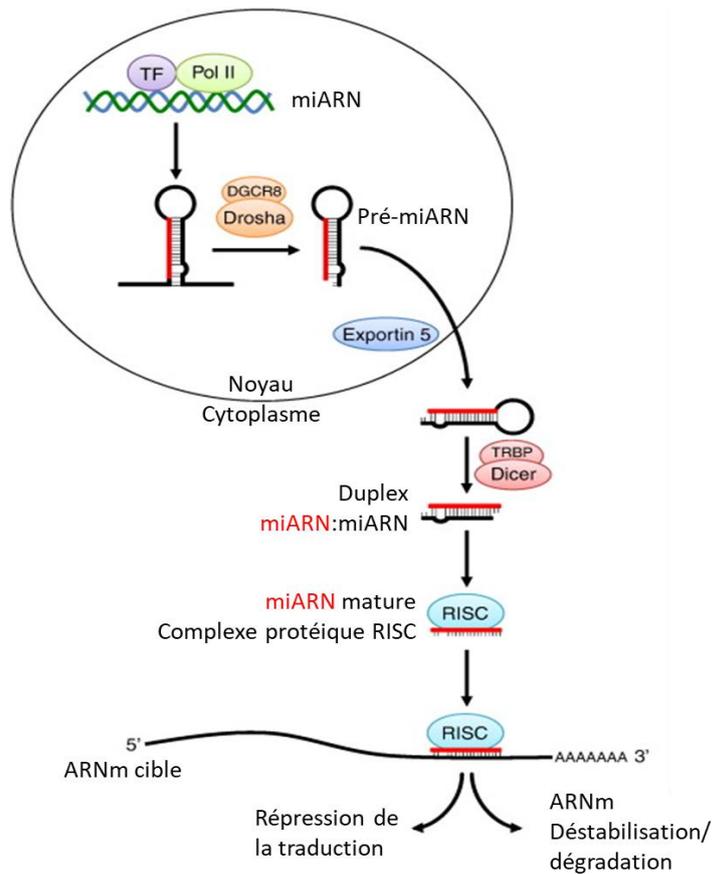


Figure 27. Schéma représentant le mode de fonctionnement des miARN dans la régulation d'un gène cible.

Adaptée de Gurtan AM et al., 2013 (164).

Les ARNnc, quant à eux, sont des ARN simple brin de plus de 200 nucléotides. Ils peuvent influencer l'expression génétique suivant différents mécanismes, en agissant directement sur l'ARNm, en modifiant la structure de la chromatine ou en interférant avec les miARN, empêchant ainsi leur action sur les gènes cibles (166). Par exemple, l'ARNnc MALAT1 inhibe le miRNA miR-34a, régulant indirectement l'expression des gènes c-Myc et Met dans les mélanomes (167). De nombreux ARNnc, dont MALAT1, ont été associés aux maladies cardiovasculaires (168). C'est le cas de MIAT (Myocardial Infarction Associated Transcript), qui est surexprimé chez des souris ayant subi un infarctus du myocarde, et qui régule l'activité du microARN miR-24 (169).

Il existe donc de nombreuses omiques qui permettent de caractériser les différents niveaux moléculaires existant dans l'organisme. Leurs études permettent d'identifier des ensembles de molécules précises, également appelées caractéristiques. L'analyse de l'expression et de la régulation des caractéristiques apporte une vision approfondie des processus biologiques existant dans des conditions d'intérêt, et permet d'étendre leur compréhension.

4.2. Les bases de données pour l'interprétation des données omiques

Dans le paragraphe précédent, nous avons discuté des différentes approches omiques existantes. Ces analyses permettent d'identifier des caractéristiques associées à une ou plusieurs conditions d'intérêt. Cependant, l'association de ces caractéristiques au phénomène étudié peut parfois être complexe à interpréter. L'augmentation des données omiques générées au cours des 30 dernières années a engendré une vaste accumulation d'informations de divers types, couvrant un large éventail de conditions, de tissus et d'organismes. La gestion méthodique et accessible de ces données est devenue une nécessité, ce qui a conduit à la création d'un grand nombre de bases de données (BDD) (figure 28). La mise à disposition de ces connaissances à la communauté scientifique s'aligne sur les principes de science ouverte et respecte les principes FAIR (Findable, Accessible, Interoperable, Reusable). Ces directives encouragent les chercheurs à rendre leurs données facilement « trouvables » en les publiant dans des bases de données « accessibles » en ligne, à les rendre « interopérables » en utilisant des formats de données compatibles, et à les rendre « réutilisables » en partageant des métadonnées détaillées.

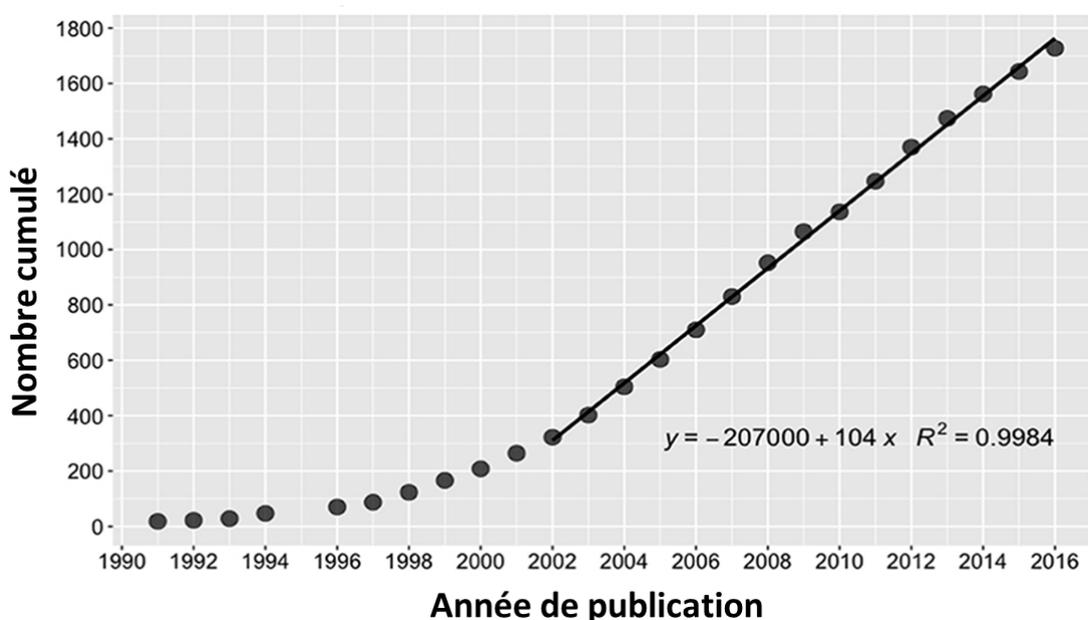


Figure 28. Évolution du nombre de nouvelles bases de données publiées dans la revue NAR (Nucleic Acid Research) entre 1991 et 2016.

Adaptée de Imker HJ (170).

Ces ressources offrent aux chercheurs un accès direct et fiable à une multitude de données, simplifiant ainsi l'analyse des caractéristiques d'intérêt. Elles permettent d'associer ces caractéristiques à des fonctions biologiques spécifiques ainsi que de connaître leurs interactions au sein de processus particuliers. Nous présenterons ici deux catégories de bases de données utiles dans la compréhension d'un processus biologique : d'une part, les BDD renfermant des informations concernant les fonctions des caractéristiques identifiées, et d'autre part, les BDD comprenant des renseignements sur les interactions entre ces caractéristiques ([tableau 2](#)).

| Type | Base de données | Features | | | | | | | | Type d'analyse | | Source | |
|-------------|-----------------|----------|----------|----------------|--------------------------|-------|---------|-------------|------------------|----------------|-----------|-------------|-------------------|
| | | Gène | Protéine | Poly-morphisme | Facteur de transcription | miARN | ARNlinc | Méthylation | Composé chimique | Expérimental | In silico | Littérature | Données publiques |
| Fonction | Gene Ontology | ■ | | | ■ | | | | | ■ | ■ | ■ | |
| | KEGG | ■ | ■ | | ■ | | | | ■ | ■ | ■ | ■ | |
| | OMIM | ■ | | ■ | ■ | | | | | ■ | ■ | ■ | |
| | DisGeNET | ■ | | | | | | | | ■ | ■ | ■ | |
| | dbSNP | | | ■ | | | | | | | ■ | ■ | |
| | HMDD | | | | | ■ | | | | ■ | ■ | ■ | |
| | lncRNAisease | | | | | | ■ | | | ■ | ■ | ■ | |
| | DiseaseMeth | | | | | | | ■ | | ■ | ■ | ■ | |
| Interaction | STRING | | ■ | | | | | | | ■ | ■ | ■ | |
| | MINT | | ■ | | | | | | | ■ | ■ | ■ | |
| | BioGRID | ■ | | | | | | | ■ | ■ | ■ | ■ | |
| | IntAct | ■ | ■ | | | | | | | ■ | ■ | ■ | |
| | GRNdb | ■ | | | ■ | | | | | ■ | ■ | ■ | ■ |
| | Coexpedia | ■ | | | | | | | | ■ | ■ | ■ | ■ |
| | ReMAP | | | | ■ | | | | | | ■ | ■ | ■ |
| | miRTarBase | ■ | | | ■ | ■ | | | | ■ | ■ | ■ | |
| | TarBase | ■ | | | ■ | ■ | | | | ■ | ■ | ■ | |
| | miRecords | ■ | | | ■ | ■ | | | | ■ | ■ | ■ | |
| | lncTarD | ■ | | | ■ | | ■ | | | ■ | ■ | ■ | |

Tableau 2. Synthèse des différentes bases de données présentées ci-dessous, en fonction de leur type, les caractéristiques contenues dedans ou l'origine des données.

4.2.1. Les bases de données de fonctions

Lors de l'identification d'un set de gènes, il est standard de chercher leurs fonctions en effectuant un enrichissement fonctionnel. Les bases de données les plus utilisées dans ce but sont Gene Ontology (GO) et KEGG Pathway (KEGG). Avec plus de 43.000 termes spécifiques à un processus biologique, appelés termes GO, la base de données GO est la plus fournie en termes de connaissances sur la fonction des gènes (171). Les GO terms sont classés en trois groupes : processus biologique (Biological Process, $n > 27.000$), fonction moléculaire (Molecular Function, $n > 11.000$) et composant cellulaire (Cellular Component, $n > 4.000$). Les associations entre les gènes et les fonctions proviennent de la littérature (observés ou prédites), et une partie d'entre elles sont sélectionnées manuellement (en cours de révisions). Les termes GO sont arrangés en arbres, avec des termes "parents" et des termes "enfants", permettant d'avoir des informations plus ou moins spécifiques sur la/les fonction(s) associée(s) à un ou plusieurs gènes. La base de données KEGG (Kyoto Encyclopedia of Gene and Genomes) (172) diffère de GO en rassemblant les termes en pathways (réactions moléculaires impliquées dans un processus biologique spécifique). Ces derniers sont organisés sous forme de réseau orienté, illustrant les interactions entre molécules. Les pathways KEGG sont divisés en de nombreuses catégories, telles que le métabolisme, les processus cellulaires, les maladies humaines ou encore le développement de médicaments. L'utilisation de ces deux bases de données permet donc d'avoir une approximation des processus et fonctions associés à un set de gènes ou de protéines d'intérêt.

Les analyses de jeux de données omiques ne se limitent pas à l'obtention de listes de gènes ou de protéines uniquement. Par exemple, dans le domaine des analyses génomiques, la base de données OMIM (Online Mendelian Inheritance in Man) (173) rassemble des gènes dont les variants ont été associés à plus de 7 000 phénotypes distincts. De manière similaire, la base de données DisGeNET recense des gènes et des variants spécifiques liés à des maladies humaines (174). dbSNP, quant à elle, se concentre uniquement sur les variants en fournissant leur position, leur(s) allèle(s), et leur conséquence(s) fonctionnelle(s), complétant les informations obtenues avec les BDD précédentes (175). Pour les ARN non codants, la base de données HMDD (Human MicroRNA Disease Database) (176) référence plus de 500 miARN expérimentalement associés à 378 maladies humaines, tandis que lncRNAdisease (177) recueille les ARNlnc et les ARNcirc expérimentalement liés à des pathologies humaines. En épigénétique, la base de données DiseaseMeth (178) recense de nombreux sites de méthylation associés à des maladies humaines.

Les bases de données mentionnées précédemment ne représentent qu'une sélection limitée parmi de nombreuses autres existantes. Toutefois, ce paragraphe illustre la diversité et la richesse des bases de données existantes, qui permettent de relier différentes données omiques à divers processus. Alors que GO et KEGG apportent des informations générales sur les processus biologiques dans lesquels sont impliqués un set de caractéristiques, OMIM, DisGeNET, HMDD, lncRNAdiseases et DiseaseMeth se spécialisent dans l'étude de pathologies spécifiques. Enfin, GeneCards et dbSNP se focalisent sur l'analyse d'une caractéristique d'intérêt, approfondissant ainsi la compréhension de cette dernière. L'utilisation de ces ressources facilite donc l'analyse et l'interprétation des données omiques.

4.2.2. Les bases de données d'interactions

Au-delà de la connaissance des fonctions ou processus biologiques impliquant un ensemble de caractéristiques, il est souvent pertinent d'analyser les interactions entre ces caractéristiques afin de mieux comprendre leur implication dans une condition d'intérêt. De nombreuses bases de données contiennent des informations sur les interacteurs, directs ou indirects, de gènes, protéines ou autres éléments génétiques. STRING (179) est une base de données ayant pour vocation de contenir toutes les interactions protéines-protéines connues, qu'elles soient prédites ou prouvées expérimentalement. Elle contient aujourd'hui plus de 20 milliards d'interactions, avec un total de 59,3 millions de protéines provenant de 12.535 organismes différents. MINT a le même objectif, en se concentrant sur les interactions entre protéines identifiées expérimentalement. Elle en répertorie plus de 136.000 dans 671 organismes (180). BioGRID (181), quant à elle, étend son catalogue de caractéristiques à des gènes, composés chimiques et modifications post-traductionnelles, avec plus de 2 millions d'interactions recensées. IntAct (182) contient elle aussi des interactions au niveau génétique et protéique, mais met l'accent sur le contexte physiologique dans lequel ces interactions ont été identifiées (cancer, tissu cardiaque...).

Les bases de données GRNdb (183) et Coexpedia (184) se distinguent des bases de données décrites précédemment par leur origine, qui n'est pas directement liée à la curation de littérature. GRNdb compile des prédictions d'interactions entre facteurs de transcription et gènes cibles à partir de jeux de données bulk et single-cell publiques. De son côté, Coexpedia rassemble des interactions prédites entre gènes à partir de données de transcriptomique (puces à ADN), également publiques.

Ces deux bases de données offrent ainsi un ensemble d'interactions prédites par leurs propres outils, en se basant sur des jeux de données publics et dans diverses conditions spécifiques. La base de données ReMAP (185) apporte des informations complémentaires en contenant des emplacements de sites de fixation de facteurs de transcription, issues de données ChIP-seq publiques.

En ce qui concerne les ARNs non codants, 3 bases de données contiennent des interactions prouvées expérimentalement entre des miARN et des gènes cibles : miRTarBase (186), TarBase (187) et miRecords (188). Du côté des ARNlnc, LncTarD (189) rassemble des interactions entre ARNlnc et gènes cibles provenant de données bulk et single-cell.

Ainsi, toutes ces bases de données offrent la possibilité d'approfondir les analyses omiques en prédisant les fonctions et les interactions des caractéristiques d'intérêt, qu'il s'agisse de gènes, protéines, ARNs non codants ou autre élément génétique/épigénétique. Le phénotype, le tissu ou le type cellulaire dans lesquelles ces associations ont été identifiées sont également renseignés. Il est toutefois indispensable de valider les informations prédites afin d'établir avec certitude la participation des caractéristiques dans une condition spécifique.

4.3. L'utilisation des omiques dans la compréhension des Cardiomyopathies Chagasiques Chroniques

4.3.1. Découverte d'une contribution génétique aux cardiomyopathies chagasiques chroniques

L'hypothèse d'une contribution génétique à la maladie de Chagas a commencé en 1945, où Hauschla TS. s'est penché sur l'impact du sexe dans la maladie de Chagas en infectant des souris avec 3 sources différentes de *T. cruzi* (190). Cette analyse a révélé un taux de répllication parasitaire plus bas chez les femelles, ce qui indique une meilleure survie chez ce groupe. En 1990, Zicker et al. ont mené une analyse sur 592 patients atteints de la maladie de Chagas (345 ASY et 247 CCC) et 529 contrôles sains. Ils ont constaté que, dans les patients infectés, les hommes avaient un risque plus grand de développer des complications cardiaques. De plus, le temps de résidence dans les zones endémiques ne semble pas influencer sur le développement des CCC. En revanche, les patients ayant des antécédents familiaux de maladies cardiaques sont plus susceptibles de développer une cardiomyopathie chagastique, suggérant une composante génétique dans cette pathologie (191).

Au début des années 1980, des études ont mis en avant une association entre la mutation de loci chez la souris et leur parasitémie après infection par *T. cruzi*, mais sans identifier de variant spécifique (192–194). À partir des années 2000, plusieurs variants impliqués dans la susceptibilité de développer des CCC (par rapport à des patients ASY) ont été identifiés, notamment dans les gènes HLA, et particulièrement dans les gènes HLA-C (195), HLA-DRB1 (196,197) et HLA-G (198). En parallèle, de nombreux gènes, majoritairement impliqués dans la réponse immunitaire, présentent des polymorphismes, qui peuvent soit prédisposer au développement des CCC, soit au contraire offrir une protection (**tableau 3**). L'ensemble des mutations présentées ici ont été identifiées de manière expérimentale (PCR puis séquençage de l'ADN), en explorant des gènes cibles d'intérêt.

| Variants prédisposant | | Variants protecteur | |
|-----------------------|-----------------------|------------------------|---------------------|
| Nom du gène | Fonction biologique | Nom du gène | Fonction biologique |
| BAT1 (199) | Réponse immunitaire | CCL2 (200) | Réponse immunitaire |
| CCL2 (201) | | CCL5 (202) | |
| COLEC11 (203) | | CCR5 (201,204) | |
| CR1 (205) | | CTLA4 (206) | |
| IL10 (207,208) | | IL12B (209) | |
| IL1B (210) | | IL17A (211) | |
| IL12B (209,212) | | IL18 (213) | |
| IL1RN (214) | | LTA (215) | |
| LTA (215) | | MBL (216) | |
| MASP2 (203) | | TIRAP (217) | |
| NFKBIL1 (218) | | | |
| PIK3 (219) | | | |
| TGFB1 (220) | | | |
| TIRAP (201) | | | |
| TNF (202,221) | | | |
| VDR (222) | | | |
| ACTC1 (223) | | Contraction musculaire | |
| FCN2 (224) | Fibrose/cicatrisation | | |

Tableau 3. Liste des variants identifiés comme prédisposant aux CCC par des méthodes expérimentales (PCR ou TaqMan), ou au contraire protecteurs.

Il est intéressant de noter que plusieurs variants situés dans un même gène peuvent avoir des effets antagonistes, comme dans le gène IL12B (209), où les SNP rs2546893 et rs919766 sont associés aux formes asymptomatiques, alors que les SNP rs919766C, 12rs2546893G et rs3024496C augmentent la susceptibilité de développer des CCC. En plus des gènes liés à l'immunité, des variants prédisposant au développement des CCC ont été identifiés dans des gènes participant à d'autres processus biologiques, comme ACTC1 (223), un gène impliqué dans la contraction cardiaque.

Par ailleurs, un nombre plus restreint de mutations a été spécifiquement associé au risque de développer des CCC sévères, comparé aux CCC modérés (**tableau 4**). Le gène CCR5, présentant deux variations associées comme protectrices des CCC (rs1799987 (204) et rs1800024 (201)), possède également deux polymorphismes particulièrement impliqués dans les formes sévères comparées aux formes modérées (rs1799988 (225) et rs3176763 (226)). Ces résultats démontrent de l'importance de ce gène dans les CCC, différentes variations affectant fortement le phénotype des individus atteints de la maladie de Chagas. A l'inverse, des mutations protègent à la fois des CCC et de l'aggravation de la maladie, telles que rs360719 et rs2043055 présentent dans le gène IL18, illustrant là encore l'importance de ce gène dans les CCC.

| Variants prédisposant aux CCC sévères | | Variants protecteur des CCC sévères | |
|---------------------------------------|---|-------------------------------------|---------------------|
| Nom du gène | Fonction biologique | Nom du gène | Fonction biologique |
| CASP1 (227) | Réponse immunitaire | CXCL10 (225) | Réponse immunitaire |
| CCR2 (226) | | CXCL9 (225) | |
| CCR5 (225,226) | | IL17A (228) | |
| IL17F (228) | | IL18 (229) | |
| NLRP1 (230) | Mort cellulaire programmée | | |
| SCN5A (231) | Muscle cardiaque, homéostasie du sodium | | |

Tableau 4. Liste des variants identifiés comme prédisposant aux CCC sévères par des méthodes expérimentales (PCR ou TaqMan), ou au contraire protecteurs.

Les méthodes d'analyse à grande échelle ont permis une exploration plus approfondie du génome, sans se focaliser sur des gènes a priori. Ainsi, en 2013, 2021 et 2022, trois analyses utilisant des données de GWAS ont été réalisées. La première étude, comparant 101 patients brésiliens atteints de CCC à 499 ASY, n'a identifié aucun variant significativement associé à la pathologie (232).

La deuxième étude a effectué une méta-analyse, combinant plusieurs analyses GWAS. Elle a ainsi été réalisée sur une plus grande cohorte, incluant les données des patients brésiliens avec 1.796 patients séropositifs (1.022 ASY, 774 CCC) et 1.104 séronégatifs provenant de 3 pays différents (Colombie, Argentine et Bolivie). Bien qu'aucune mutation n'ait été associée à la maladie de Chagas, un variant situé à proximité du gène SAC3D1, participant à la régulation de la réponse immunitaire, semble impliqué dans le développement des formes cardiaques (233). La dernière étude s'est restreinte sur la population brésilienne, incluant 2.964 patients (581 ASY, 2.383 CCC). Elle a mis en évidence un variant associé à la réponse immunitaire, situé dans le gène AKAIN1 (C18orf42) (234).

Compte tenu des observations précédentes au sein de familles ayant des antécédents de maladies cardiaques, une analyse Whole Exome Sequencing a été effectuée en 2022 dans le but de détecter des variants rares. Elle a mis en évidence 321 mutations identifiées dans 6 familles présentant à la fois des cas de CCC et des individus asymptomatiques (235). Ces variations génétiques spécifiques aux CCC sont absentes chez les individus asymptomatiques présents dans les 6 familles, mais également chez des individus asymptomatiques non apparentés. Parmi ces mutations, 87 sont prédites comme étant pathogéniques, et 22 ont été confirmées expérimentalement. Elles sont localisées dans 20 nouveaux gènes jamais identifiés comme porteurs de polymorphismes associés aux CCC, et essentiellement liés à des processus inflammatoires et mitochondriaux.

L'ensemble de ces résultats tend à confirmer la présence d'une composante génétique prédisposant au développement des CCC, et parfois spécifiquement à leur sévérité. L'importance du fond génétique dans la sévérité de la maladie a également été observée chez la souris, appuyant davantage ces observations (236). Les mutations détectées peuvent varier d'une population à l'autre, voire d'une famille à l'autre, mais la plupart se retrouvent dans des gènes associés aux mêmes processus biologiques (en particulier l'immunité et les processus mitochondriaux). Cela met en évidence l'importance des perturbations dans ces processus biologiques dans l'évolution des CCC.

4.3.2. Identification et caractérisation des processus pathogéniques associés aux cardiomyopathies chagasiennes chroniques

La découverte de mutations a permis de confirmer la présence d'une l'implication génétique dans le développement des CCC, mais n'explique pas, ou seulement partiellement, son processus pathogénique. Afin de comprendre cette pathologie, de nombreuses études ont été mises en place.

Dès 1997, différentes cytokines ont été liées aux formes chroniques de la maladie de Chagas, telles que l'IL-5, l'IL-10, l'IL-13 et l'IFN- γ , détectées dans le sang des patients (237). Quelques années plus tard, le TNF- α a également été associé aux CCC (238), et sa concentration a été corrélée à la sévérité de la maladie (239). C'est en 2005 que Cunha Neto E. et al. ont réalisé la première étude transcriptomique à grande échelle (240). En utilisant une puce à ADN, ils ont pu analyser l'expression de 2.493 gènes connus dans du tissu cardiaque de CCC, cardiomyopathie dilatée (CMD) non chagasique et contrôle sain. Dans cette analyse, ils ont identifié 1.047 gènes différentiellement exprimés (Differentially Expressed Genes, DEG) entre les contrôles et les CCC, dont 162 gènes associés à des fonctions connues. Ils ont également mis en évidence 396 DEGs entre les contrôles et les CMD (85 DEG liés à des fonctions connues) ([figure 29A](#)). Seulement 34 gènes associés à des fonctions connues ont été trouvés en commun entre les CCC et les CMD, majoritairement impliqués dans la contraction cardiaque (actine, myosine), soulignant ainsi des dysrégulations spécifiques aux CCC ([figure 29B](#)). Les DEG CCC-spécifiques, quant à eux, ont été majoritairement assignés à la réponse immunitaire (réponse à l'IFN- γ , TCR, MHC), ainsi qu'au métabolisme de l'énergie. Ces différences ont également été relevées les années suivantes, avec la mise en avant d'une production d'ATP plus faible chez les CCC (241) et une production plus importante de gènes associés à l'immunité (CXCL10, TBX21, GATA3, l'IFN- γ , IL-17 et IL-18) (242,243) comparés aux autres types de cardiomyopathies. La participation plus importante de l'immunité dans les CCC par rapport aux autres cardiomyopathies semble triviale au regard de son caractère parasitaire. Toutefois, cette affirmation doit être modulée par le fait qu'en phase chronique, soit des années après l'infection, la parasitémie est faible (244), soulignant l'anormalité d'une telle réponse inflammatoire. Un ensemble d'études a contribué à la compréhension de la réponse immunitaire exacerbée chez les CCC. L'implication de l'IFN- γ et du TNF- α , acteurs importants dans les CCC, a été largement répertoriée au niveau génétique et protéique à la fois chez l'humain (62,65,242,245–247), la souris (248–251), le rat (252) et le chien (253). L'utilisation de techniques d'analyses à grande échelle a permis d'identifier de nombreuses caractéristiques liées à la réponse immunitaire, notamment des cytokines, impliquées dans les CCC. Ces dernières ont été associées aux récepteurs Toll-like, aux interactions entre cytokines, aux gènes HLA, aux cellules NK et, plus particulièrement, à la réponse immunitaire médiée par les lymphocytes (différenciation des lymphocytes Th1 et Th2, régulation de l'activation des lymphocytes, récepteur des cellules B et T, processus de liaison aux antigènes, activation des lymphocytes B par les antigènes...) (57,240,254–256).

Les processus immunitaires impliqués dans les CCC ne seront pas plus explorés ici, ces derniers ayant été décrits dans le chapitre 3.

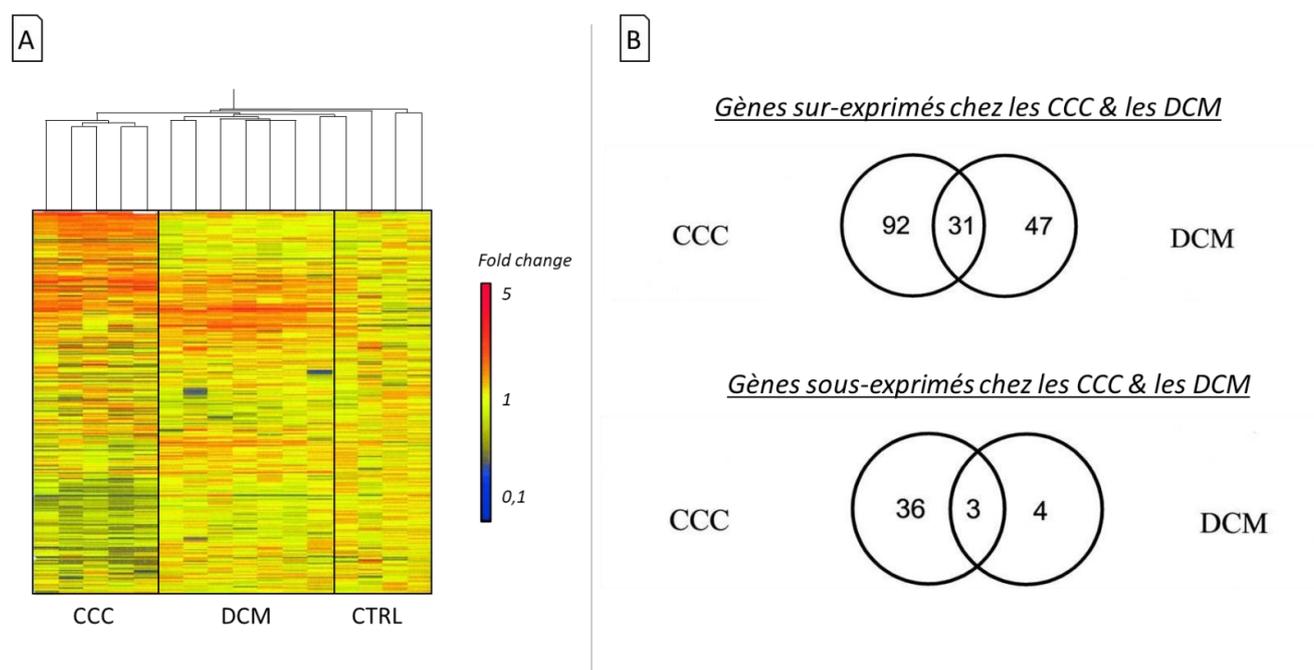


Figure 29. Comparaison des profils d’expression entre les contrôles sains, les CCC et les CMD.

A) Clustering hiérarchique des individus en fonction de leur phénotype basé sur l’expression des 10.368 gènes analysés. B) Diagrammes de Venn comparant les gènes sur- et sous-exprimés chez les CCC et les CMD associés à des fonctions biologiques d’intérêt.

Adapté de Cunha Neto E. et al., 2005 (240)

Bien que la majorité des caractéristiques dérégulées dans les CCC soient associées à la réponse immunitaire, d’autres processus biologiques semblent impliqués dans cette pathologie. L’implication de dysfonctionnements mitochondriaux dans des cardiomyopathies est un phénomène connu (257), appuyé par la découverte de mutations mitochondriales chez les CCC (235). De nombreux gènes dérégulés associés à la phosphorylation oxydative, au stress oxydatif, à la chaîne de transfert d’électrons ou à la production d’ATP ont été mis en évidence chez l’humain (137,240,255,258,259) et la souris (249). De plus, des gènes impliqués dans la contraction musculaire et le fonctionnement cardiaque ont été relevés, notamment associés à l’actine (ACTC1, ACTN2, ACTN3) et à la myosine (MYH6, MYH7, MYH11, MYH14, MYL2, MYL3, MYL4, MYL7) (223,240,254,259–261). Enfin, une atteinte du système nerveux avec une destruction des neurones a été relevée (262), potentiellement liée à la production d’adipocytokines (263).

4.3.3. Recherche des causes de dérégulations de Chagas

L'identification des principaux processus biologiques associés aux CCC a permis une première étape dans la compréhension de cette pathologie. Toutefois, les causes de ces dérégulations sont encore méconnues, et restent à être élucidées. Dans cet objectif, la recherche de mutations pouvant influencer sur l'expression génétique ou la fonctionnalité protéique est essentielle. Par exemple, la mutation rs1800896 présente dans le promoteur du gène IL-10 pourrait être associée à son expression dans les CCC, diminuant avec la sévérité de la maladie (207). De même, les mutations rs10336 (CXL9) et rs3921 (CXCL10), protégeant contre les formes sévères des CCC, sont associées avec une diminution de l'expression de cytokines, telles que CCL3, CCL5, CXCL9, CXCL10, CXCR3, CCL17 et CCR4 (225). Parfois, le lien entre variation et expression génétique est moins évident, comme dans le cas du gène MBL2. En effet, l'impact de mutations dans le 1er exon de ce gène conduit à une diminution de sa production, comme cela a déjà été montré dans des maladies infectieuses, inflammatoires ou auto-immunes (216). Aucune dérégulation de ce gène n'a été trouvée chez les CCC, mais une de ses mutations a été associée à une baisse d'IL-9 et de CCL5.

Au-delà de l'aspect génétique, de nombreuses variations épigénétiques ont été identifiées. Laugier L. et al. ont mené une étude combinant données d'expression et de méthylation sur du tissu cardiaque de CCC et de contrôle sain (260). Ils ont identifié 7.595 CpG (localisés dans 4.720 gènes) et 1.535 gènes dérégulés, dont 399 en commun. Parmi ces gènes, 23 sont impliqués dans la réponse immunitaire ou les fonctions cardiaques et contiennent au moins 5 CpG différenciellement méthylés (DMP, Differentially Methylated Position). 12 d'entre eux possèdent différents DMP dans leur promoteur. Des analyses in vitro menées sur des cellules cardiaques ont montré que la méthylation du promoteur de KCNA4, KCNIP4 (canaux potassiques), SMOC2 (cicatrisation et remodelage de la matrice extra-cellulaire), RUNX3, CD6, HLA-DPA1 et TRAF3IP3 (réponse immunitaire) module leur expression. 386 DEG ayant des DMP présentaient une corrélation entre leur niveau d'expression et de méthylation. En restreignant l'étude aux gènes présentant des CpG dans leur promoteur, les auteurs ont identifié plusieurs sites de fixation de facteurs de transcription pouvant être affectés, notamment dans 10 gènes ciblés par Sp1, incluant KCNA4, KCNIP4, SMOC2, RUNX3 et HLA-DPA1. Le facteur de transcription TBX21 (T-bet), spécifique des lymphocytes Th1, a également été associé aux CCC (255). RUNX3 et TBX21 interagissent dans les lymphocytes T pour promouvoir la production d'IFN- γ , et pourraient donc être fortement impliqués dans le processus pathogénique. Par ailleurs, 5 mutations ont été associées à des sites de méthylation chez des patients CCC par rapport à des ASY.

Parmi celles-ci, deux mutations ont été trouvées dans le promoteur du gène SPDYC, tandis qu'une autre a été identifiée dans le promoteur du gène TM7SF2, ces deux gènes étant impliqués dans le métabolisme.

De plus, 6 SNP situés à proximité de la région promotrice du gène SAC3D1 présentait des corrélations significatives avec des sites de méthylation à la fois chez les ASY et les CCC (264). Des études antérieures ont déjà établi une association entre une mutation proche de ce gène et le développement des CCC (233).

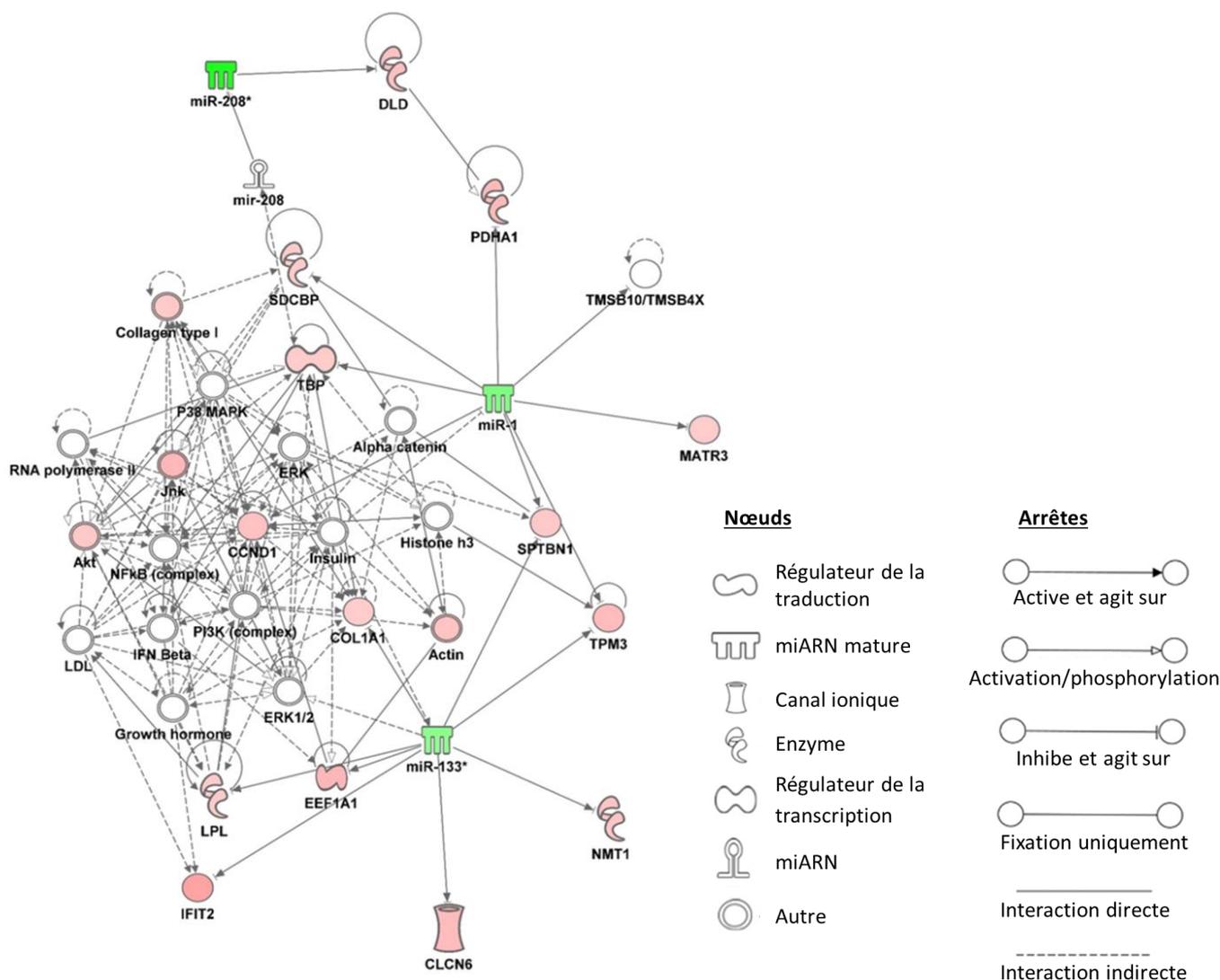
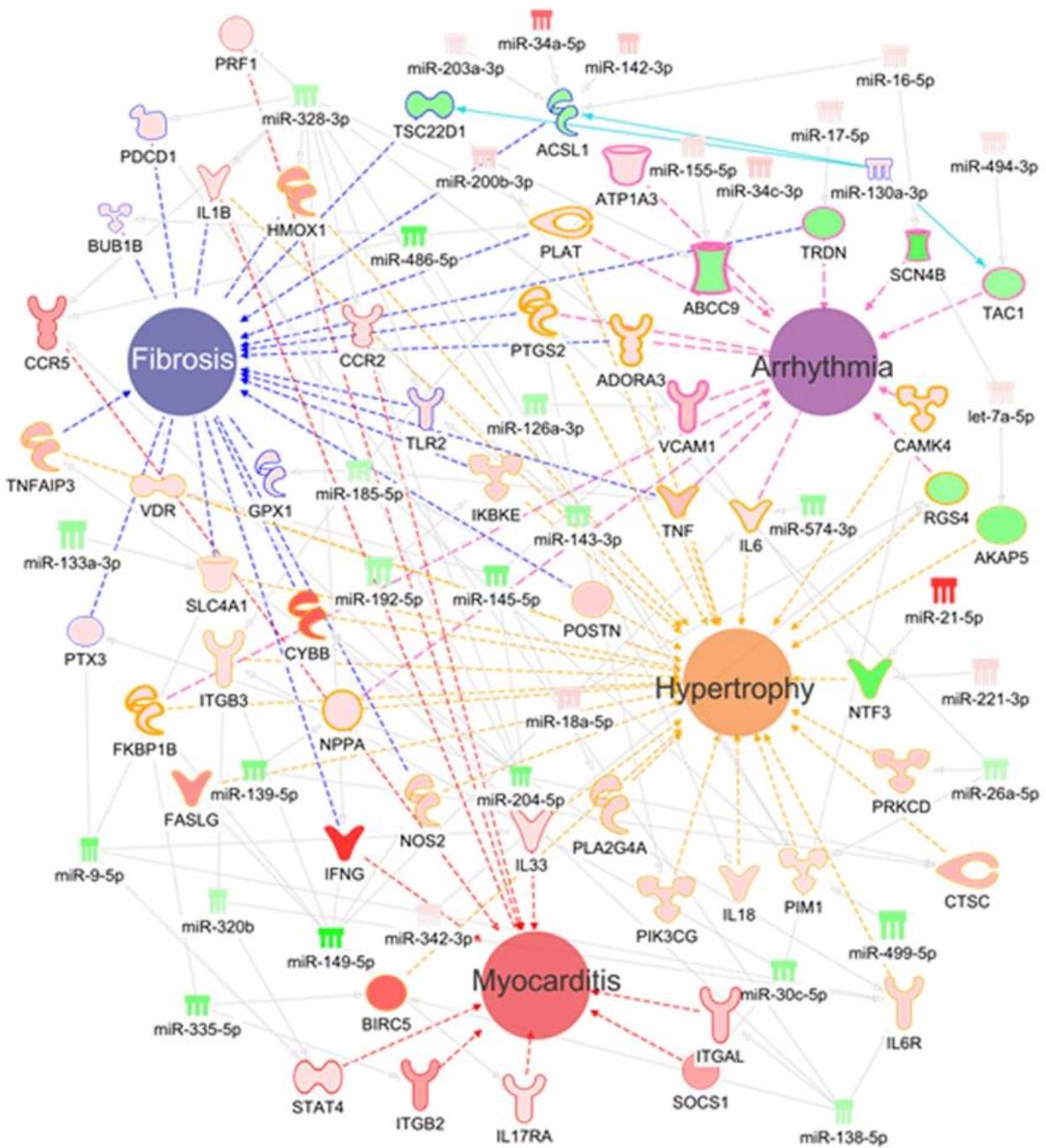


Figure 30. Réseau d'interaction entre les miARN dérégulés chez les CCC et leurs cibles potentielles.

Adaptée de Ferreira LRP et al., 2014 (68).

De nombreux ARN non codants sont également impliqués dans cette pathologie, notamment des miARN. Une étude comparant les profils miRNomiques du tissu cardiaque de patients atteints de CCC, de CMD ou de contrôles sains a mis en évidence l'importance de miR-133b et miR-208a dans les cardiomyopathies, et miR-1, miR-133a, et miR-208b dans les cardiomyopathies chagasiques spécifiquement (68). Tous sont sous-exprimés dans les pathologies cardiaques. Douze gènes cibles de ces miARN ont été précédemment associés aux CCC (240). Un réseau d'interactions prédictif a pu être établi avec le logiciel IPA, intégrant les données transcriptomiques et miRNomiques, ainsi que des informations connues dans les bases de données publiques. Les composants de ce réseau sont associés aux infarctus du myocarde, incluant la voie NF-kB (**figure 30**). De plus, les miARN miR-19a-3p, miR-21-5p et miR-29b-3p ont été associés au développement des CCC, ayant une expression différentielle dans le plasma de patients atteints de CCC comparés à des contrôles sains et asymptomatiques (265). Ces derniers sont également surexprimés dans le tissu cardiaque de CCC dilatée par rapport à des contrôles sains. Dans ce contexte cardiaque, miR-30a-5p, miR-199b-5p et miR-208a-3p ont également été identifiés. Tous ces microARN sont déjà connus pour être impliqués dans des cardiomyopathies.

Dans le modèle murin, 113 miARN sont dérégulés lors de l'infection par *T. cruzi* à trois moments différents (15, 30 et 45 jours post-infection) par rapport aux souris témoins non infectées. Plus particulièrement, le nombre de miARN dont l'expression est perturbée augmente au fil du temps après, passant de 19 à 96. Parmi eux, 17 sont dérégulés tout au long de l'infection, dont 9 présentant une corrélation avec le niveau de parasitémie et le temps de contraction cardiaque (intervalle QT). Comme précédemment, l'utilisation de la base de données IPA Ingenuity a permis de prédire les cibles de ces 9 miARN, incluant le gène KCNA1 (266). Une deuxième étude a combiné ces résultats avec une étude transcriptomique réalisée aux mêmes temps post-infection. Sur les 1.685 DEG identifiés, 848 sont prédits pour être ciblés par au moins un des 113 miARN. Ces derniers sont notamment impliqués dans la présentation d'antigène, l'apoptose et les processus liés aux lymphocytes Th1 (267). Un réseau d'interactions prédites avec le logiciel IPA a mis en avant l'association entre les miARN-gènes cibles et différentes altérations cardiaques (fibrose, arythmie, hypertrophie et myocardite) (**figure 31**).



Récepteur transmembranaire
 Kinase
 miARN mature
 Transporteur
 Phosphatase
 Autre
 Enzyme
 Peptidase
 Régulateur transcriptionnel
 Canal ionique
 Cytokine/facteur de croissance

Figure 31. Réseau d'interactions prédites in silico entre les miARN et les gènes dérégulés chez la souris 30 jours post-infection par *T. cruzi*.

Adaptée de Ludmila RPF et al., 2017 (267).

Enfin, l'absence du miR-155 a également été associée à l'infection cardiaque du cœur de souris par *T. cruzi*, et à la baisse de production d'IFN- γ et de TNF- α dans le sang (268). Il est par ailleurs plus exprimé dans le tissu cardiaque de CCC comparé à des CMD et des contrôles, bien que cette différence ne soit pas significative (68).

Les miARN ne sont pas les seuls ARN non codants identifiés chez les CCC. En effet, l'ARN long non-codant MIAT (Myocardial Infarction-Associated Transcript) est sur-exprimé dans le cœur de patients CCC, comparé à des contrôles sains et des individus atteints de cardiomyopathie dilatée. Son implication dans les CCC a été confirmée chez la souris, où il est faiblement exprimé, voire absent du tissu cardiaque des contrôles sains, et fortement exprimé après infection par *T. cruzi*. L'expression de MIAT a été corrélée avec la durée de l'intervalle QT, et sa production maximale a été observée 30 jours post-infection, coïncidant avec l'apparition des premières anomalies sur l'ECG. En revanche, il n'existe aucune corrélation entre MIAT et le niveau de parasitémie, suggérant son implication dans les mécanismes cardiaques plutôt qu'immunitaires (269). Il est à noter que des niveaux élevés d'expression de MIAT, ainsi que des variations génétiques dans son gène, ont déjà été associés aux risques d'infarctus du myocarde.

L'ensemble de ces analyses, effectuées à travers différentes analyses omiques, contribuent à la compréhension du processus pathogénique dans les CCC. Elles ont mis en avant de nombreuses caractéristiques impliquées dans les CCC, à plusieurs niveaux moléculaires, soulignant ainsi la complexité des systèmes biologiques. À ce titre, quelques études ont établi des liens entre l'expression génique et des polymorphismes, sites de méthylation ou miARN grâce à des connaissances publiques. Elles ont ainsi illustré différents niveaux de régulation, chacun expliquant partiellement les résultats observés. L'intégration de différents niveaux moléculaires au sein d'une analyse est essentielle pour obtenir une compréhension approfondie des processus biologiques. Bien que des investigations supplémentaires soient requises pour acquérir une compréhension totale des CCC, ces découvertes marquent néanmoins un point de départ significatif dans cette direction.

Chapitre 5 : Les analyses multi-omiques

Les chapitres précédents ont présenté la maladie de Chagas, le processus pathogénique apparaissant dans les formes cardiaques et les différentes analyses omiques employées pour en approfondir la compréhension. Dans cette partie, la notion d'intégrations multi-omiques sera abordée, détaillant leurs intérêts ainsi que leurs limites.

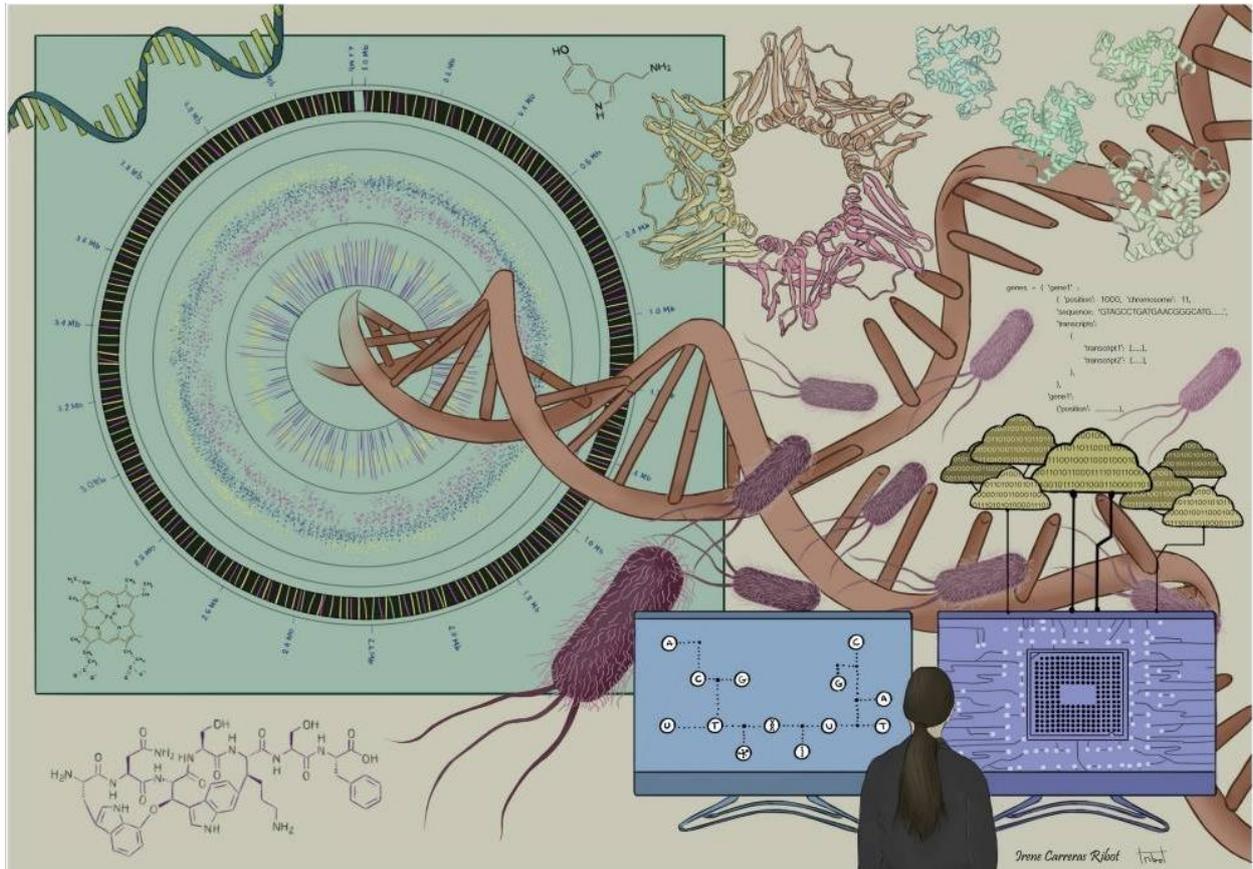


Figure 32. Représentation artistique de la complexité des analyses multi-omiques.

Adaptée de Krassowski M et al., 2020 (270).

5.1. Introduction aux analyses multi-omiques

L'intégration multi-omiques, est une méthode d'analyse qui combine plusieurs types d'omiques afin d'étudier des processus biologiques complexes, identifier des biomarqueurs ou différencier plusieurs sous-groupes d'échantillons.

Avec l'avancée des techniques d'analyses biologiques à grande échelle, de nombreuses données spécifiques et complémentaires sont générées. Les méthodes multi-omiques cherchent à combiner ces informations afin d'obtenir des résultats plus complets et précis (figure 32).

5.1.1. Les débuts du multi-omiques

Le terme "multiomic" apparaît pour la première fois dans une revue publiée en 2005, recensant différentes méthodes pour inférer des réseaux biologiques à partir de données d'expression (majoritairement génétiques issues de puces à ADN) (271). Les premières intégrations multi-omiques se sont concentrées sur l'étude conjointe de deux omiques. Par exemple, dans le cadre d'une recherche portant sur le carcinome hépatocellulaire, il a été observé que 40 gènes et protéines (sur un total de 164 caractéristiques étudiées) présentaient des variations d'expression entre les patients et les contrôles sains. Toutefois, parmi ces 40, seules 16 montraient une similitude de tendance (272). De nombreuses études ont également montré l'impact de mutations spécifiques sur l'expression génétique, comme cela a été réalisé dans le cadre de la maladie de Chagas (207,225). L'utilisation de bases de données publiques a permis d'approfondir ces analyses, apportant des connaissances sur toutes les caractéristiques connues pour interagir ensemble. Ce type d'approches a également été utilisé dans le cadre de la maladie de Chagas, notamment pour identifier les cibles de miARN associés aux CCC (68).

En parallèle, des analyses de corrélation ont été menées afin d'associer l'expression de plusieurs caractéristiques. Une étude portant sur les gliomes a identifié des corrélations entre le nombre de copies de segments d'ADN et l'expression de gènes spécifiques, mettant en évidence des candidats prometteurs pour le développement de thérapies dans ces types de cancer (273). Chez les CCC, l'impact de CpG sur l'expression des gènes associés a pu être prédit via des analyses de corrélations, et confirmé expérimentalement (255).

Ces différentes approches ont confirmé l'intérêt des analyses conjointes de différentes omiques dans la compréhension de phénotypes d'intérêt. En utilisant des données d'expression ou des connaissances publiques, elles ont révélé l'association de caractéristiques provenant de deux omiques différentes, impliquée dans le développement de pathologies.

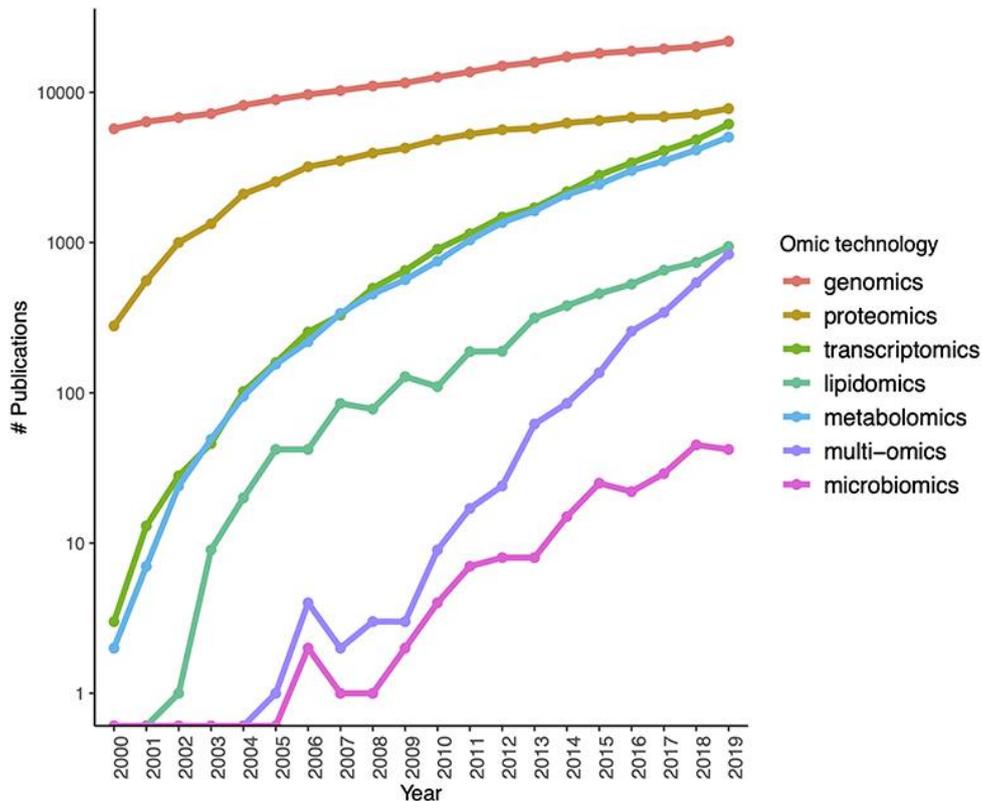


Figure 33. Evolution du nombre de publications recensées sur Pubmed incluant le terme “multiomic”.

Adaptée de Anda-Jáuregui et al., 2020 (274).

Depuis les années 2010, de très nombreuses méthodes et études multi-omiques apparaissent (figure 33) et ce nombre ne cesse de croître (3.098 publications contenant le terme “multiomic” publiées en 2022). Krassowski et al. ont défini le multi-omiques comme « trois ensembles de données omiques ou plus provenant de différentes couches de la régulation biologique, qui ne se limitent pas nécessairement à un niveau (exclusivement dérivées de l'acide nucléique/ADN, c'est-à-dire épigénomique, transcriptomique et génomique)» (270). Dans cette idée, des méthodes combinant différents types d'omiques ont été développées, en se servant de données publiques ou d'analyses de corrélation afin d'inférer des réseaux de régulations (275–277). Ces approches suivent la même logique que précédemment : la sélection de caractéristiques est réalisée au niveau simple omique, puis des liens entre ces caractéristiques sont établis. Dans le cadre de l'intégration multi-source, elles se rapprochent de ce que l'on appelle "l'intégration tardive" (late fusion), où la fusion des différentes omiques se produit après le traitement individuel de chacune d'elles.

5.1.2. L'intégration précoce

Simultanément, dès 2010, des méthodes d'intégration précoce ont vu le jour, fusionnant directement toutes les caractéristiques de différentes omiques sans pré-sélection. Il s'agit ici d'intégration précoce (early fusion). Par rapport à l'intégration tardive (late fusion), l'intégration précoce présente l'avantage de considérer toutes les caractéristiques des divers jeux de données. Ainsi, aucune caractéristique potentiellement intéressante n'est omise et les interrelations entre les jeux de données sont préservées (figure 34). En revanche, l'interprétabilité est plus complexe, certaines associations entre différentes caractéristiques étant biologiquement difficilement interprétables.

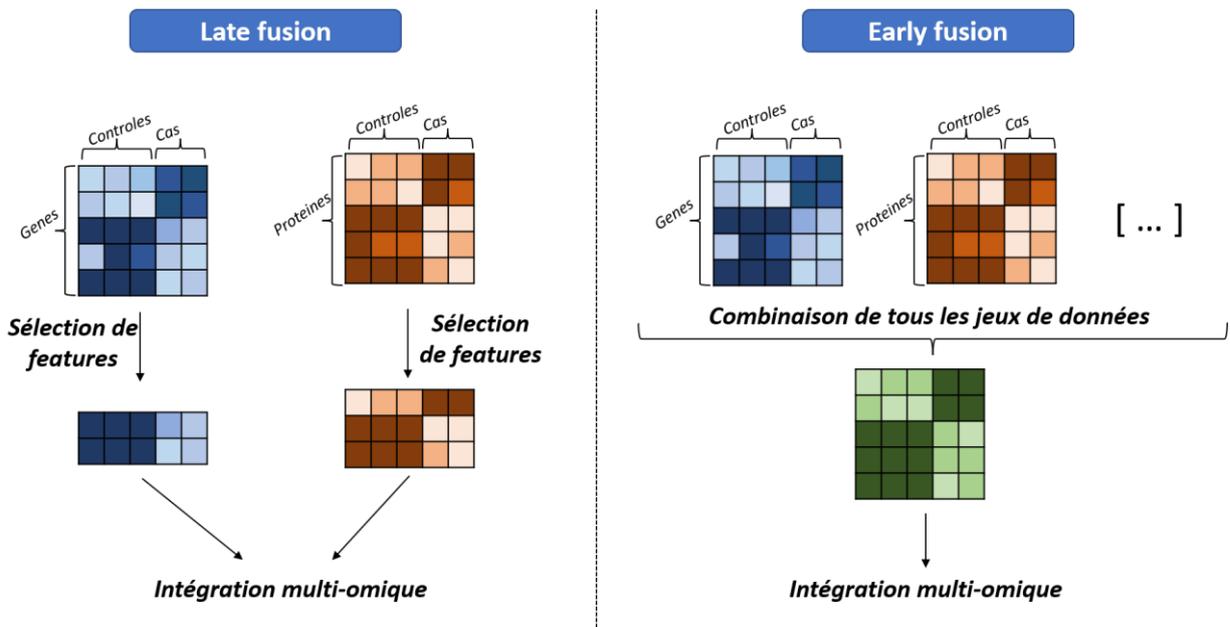


Figure 34. Représentation schématique de la différence entre l'intégration tardive (late fusion) et l'intégration précoce (early fusion) de données provenant de différentes sources.

Il existe de nombreuses méthodes d'intégration précoce, reposant sur différents principes, notamment la construction de réseaux d'interaction, la factorisation matricielle, et l'apprentissage automatique (figure 35). Cette discussion se concentrera sur celles utilisées dans l'utilisation de données omiques pour la compréhension et la caractérisation de processus biologiques.

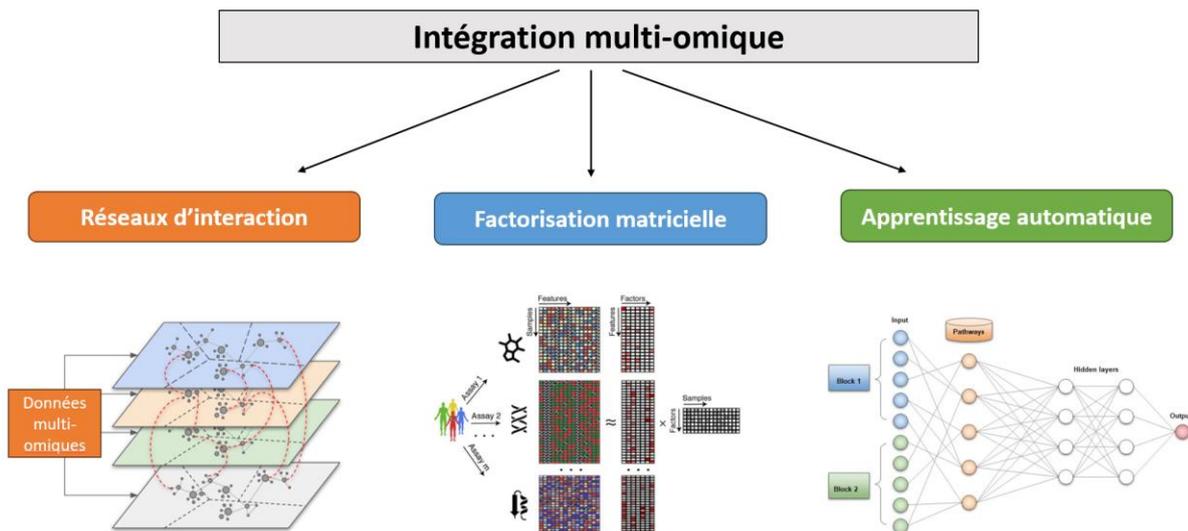


Figure 35. Illustration des trois types de méthodes détaillées ci-dessous.

Adaptée de Bodein A. et al., 2022 (278), Argelaguet R. et al., 2018 (279) et Picard M et al., 2021 (280).

- Les méthodes basées sur les réseaux

Un réseau est une représentation d'un ensemble de relations entre différentes entités. Ces entités peuvent être des caractéristiques (gènes, protéines...), des échantillons (cellules, tissus, patients...) ou d'autres types d'informations (processus biologiques, phénotypes...). L'étude d'associations de caractéristiques provenant de différents jeux de données omiques à travers un réseau d'interactions a pour avantage de permettre une interprétabilité directe des résultats. Cependant, la conception de ces réseaux peut être complexe, et déterminante pour les conclusions biologiques qui en découlent. À titre d'exemple, PARADIGM (PATHway Recognition Algorithm using Data Integration on Genomic Models) (281) crée un score d'activité en comparant le niveau d'expression des caractéristiques dans une condition spécifique à celle de contrôles. Ce score est alors utilisé dans des réseaux issus de la base de données PID (Pathway Interaction Database) (282) afin d'estimer les processus biologiques affectés dans cette condition particulière. Han Si et al. ont appliqué PARADIGM à des données de CNV et d'expression génique provenant du TCGA, leur permettant de comprendre le lien entre les complexes TP63 et NF-kB dans les cancers épidermoïdes (283). iOmisPASS est un autre outil reposant sur le même principe, mais permettant à l'utilisateur de fournir ses propres réseaux biologiques (284).

Il a été utilisé afin de comprendre les effets secondaires faisant suite à un infarctus du myocarde, notamment l'insuffisance cardiaque (285). En fusionnant des jeux de données couvrant des protéines, des lipides, des biomarqueurs cliniques courants et des mesures échocardiographiques, les auteurs ont réussi à établir des liens entre des biomarqueurs bien connus, comme le NT-proBNP, et des indices du fonctionnement cardiaque, tels que la fraction d'éjection du ventricule gauche. De plus, ils ont établi des connexions entre certaines protéines et lipides à des biomarqueurs et des mesures échocardiographiques, permettant de mieux appréhender l'origine de ces complications et de découvrir de nouveaux biomarqueurs cardiaques.

- Les méthodes basées sur la factorisation matricielle

La factorisation matricielle est un processus de réduction de la dimensionnalité, aussi appelée réduction de dimension. Il s'agit de transformer des données initiales dans un espace de grande dimension en des données dans un espace de plus petite dimension. Ces nouvelles données représentent les données d'origine. Plus particulièrement, une matrice de données (A , avec N caractéristiques et M échantillons) est décomposée en plusieurs matrices de plus petites dimensions (prenons-en deux ici à titre d'exemple, notées W et H). L'objectif est que le produit de W et H se rapproche au maximum de la matrice originale A . La matrice W a pour dimension N caractéristiques \times K facteurs, et la matrice H K facteurs \times M échantillons (**figure 36**). Les facteurs, dont le nombre est variable d'un jeu de données à l'autre, sont des représentations des caractéristiques et des échantillons qui partagent des variations similaires. L'analyse de ces facteurs est donc très intéressante que ce soit dans l'étude de groupes de caractéristiques ou de groupes d'échantillons communs, fournissant des informations complexes tout en préservant les relations entre les données.

Ces analyses peuvent être réalisées de manière non supervisée, c'est-à-dire sans recourir à des informations sur le phénotype des échantillons, comme c'est le cas avec les outils JIVE (Joint and Individual Variation Explained) (286), MCI (Multiple Co-Inertia Analysis) (287), MOFA (Multi-Omics Factor Analysis) (279,288) et DIABLO (Data Integration Analysis for Biomarker discovery using Latent cOmponents) (289). Alternativement, elles peuvent être effectuées de manière supervisée, en utilisant les caractéristiques phénotypiques des échantillons pour les distinguer, comme avec la méthode DIABLO.

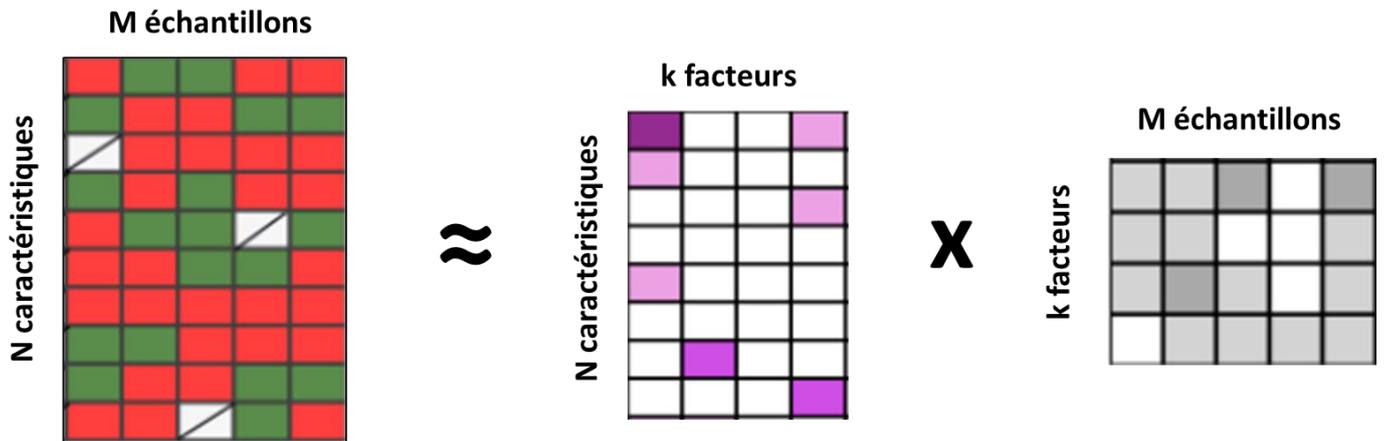


Figure 36. Représentation schématique du principe de la factorisation matricielle.

Adaptée Argelaguet R. et al., 2018 (279)

Ces différents outils ont été appliqués à de nombreux champs diversifiés de la biologie, comme la microbiologie, où de nombreuses études ont associé la production de métabolites à la constitution en souches bactérienne de microbiotes, cherchant à comprendre le lien entre hôtes et microbiome (290,291). Francavilla A. et al. ont utilisé DIABLO pour intégrer des données de miRNomique et de microbiomique dans le cadre de la maladie coeliaque (292). Ils ont ainsi mis en avant de fortes corrélations entre l'expression de miRNAs spécifiques et la présence de souches bactériennes, suggérant de nouvelles interactions hôtes-microbiome. En neurologie, Zhao Y. et al. ont confirmé l'efficacité de l'intégration de différentes mesures du cerveau dans la prédiction de l'âge cérébral, le sexe et les capacités cognitives générales des individus (293). En utilisant JIVE, ils ont trouvé des sources de variations communes à différentes zones du cerveau, représentatives du développement cortical. 3 ans plus tard, ces auteurs ont également effectué une analyse similaire afin de comprendre l'impact des écrans sur le développement du cerveau (294).

- Les méthodes basées sur l'apprentissage automatique

L'apprentissage automatique (Machine Learning, ML) est une branche de l'intelligence artificielle qui se focalise sur le développement d'algorithmes permettant aux ordinateurs d'apprendre à partir de données existantes. L'objectif est d'utiliser une fonction mathématique modélisant et s'ajustant aux données afin de prédire des informations d'intérêt ou d'étudier des structures complexes.

La factorisation matricielle peut être considérée comme de l'apprentissage automatique, mais se concentre principalement sur les relations linéaires entre les variables. De nombreuses méthodes d'apprentissage automatique permettent de modéliser des données complexes et hétérogènes en prenant en compte des combinaisons non linéaires de caractéristiques.

L'utilisation de l'apprentissage automatique dans le cadre d'intégrations multi-omiques tend à se développer de plus en plus, de nombreux algorithmes ayant déjà fait leur preuve dans le cadre de l'apprentissage multi-vue (à partir de plusieurs sources de données). En particulier, l'utilisation de réseaux de neurones (deep learning) se révèle efficace pour traiter simultanément de vastes ensembles de données, ce qui est particulièrement pertinent dans le contexte de l'intégration de données multi-source. Un autoencodeur est un type d'architecture où les neurones sont organisés en sablier : les informations sont donc compressées puis décompressées, l'information à la sortie devant être la plus proche possible de celle fournie en entrée.

De ce fait, la couche compressée, soit celle contenant le moins de neurones, contient l'information essentielle permettant de reconstituer le jeu de données initial. Certaines méthodes d'apprentissage automatique intègrent également des connaissances biologiques afin d'identifier des processus biologiques d'intérêt pendant la phase d'apprentissage. La principale limite aux approches basées sur l'apprentissage automatique est qu'elles nécessitent un grand nombre d'observations afin d'effectuer un apprentissage suffisant, là où de nombreux jeux de données multi-omiques ont un nombre d'échantillons limités. De nombreuses méthodes d'apprentissage automatique ont été développées (280,295,296), et sont majoritairement dédiées à la détection de biomarqueurs.

Malgré l'intérêt que représentent les méthodes d'intégration précoces, elles sont, à l'heure actuelle, peu appliquées. En prenant comme exemple les outils présentés ci-dessus, malgré leur nombre parfois important de citations (allant de 42 à 413), seul 14% sont employés dans des études biologiques (181/1.311), la majorité étant citées dans d'autres méthodes, ou dans des revues ([tableau 5](#)). Au-delà de la nouveauté de ce type d'approches, leur difficulté d'application et d'interprétation pourraient être en cause.

| Méthode | Date de publication | Nombre total de citations | Nombre de citations dans des méthodes ou revues | Nombre de citations dans des études biologiques |
|----------------|---------------------|---------------------------|---|---|
| PARADIGM | 2010 | 395 | 368 | 27 |
| iOmicsPASS | 2019 | 42 | 40 | 2 |
| JIVE | 2013 | 154 | 148 | 6 |
| MCIA | 2014 | 102 | 90 | 12 |
| MOFA/ MOFA+ | 2018 /2020 | 413 | 363 | 50 |
| DIABLO | 2019 | 205 | 121 | 84 |

Tableau 5. Nombre de citations (totales et par catégorie d'article) des méthodes d'intégration multi-omiques présentées ci-dessus (juillet 2023).

5.2. Les difficultés techniques liées à l'intégration multi-omiques

5.2.1. Les échantillons manquants

Le développement des analyses multi-omiques fait suite aux grandes avancées technologiques dans les méthodes d'analyses de données à grande échelle. De nombreuses études portent sur des données déjà traitées au niveau simple omiques, parfois des années plus tôt. De ce fait, ces dernières n'ont pas été initialement conçues avec pour objectif une intégration commune. Les différentes omiques peuvent donc inclure des échantillons différents, pour des raisons de budget ou d'indisponibilité des échantillons à ce moment-là. La majorité des méthodes d'intégration multi-omiques nécessite la présence des mêmes échantillons à travers les différentes omiques. La première solution est de poursuivre l'analyse avec uniquement les échantillons en commun, mais cela n'est pas toujours envisageable au regard du nombre d'échantillons par omique. La deuxième solution est d'imputer les échantillons manquants dans les omiques concernés, si ce nombre n'est pas trop important. L'utilisation de la méthode des k-plus proches voisins est couramment utilisée pour l'imputation de données manquantes dans un jeu de données simple omique. Le principe est de rechercher les plus proches voisins d'un échantillon, et d'imputer les données manquantes d'après la moyenne d'expression de ces données chez ses voisins.

La recherche des voisins reste difficile dans un jeu de données multi-omiques, compte tenu des différentes tailles des données omiques. La méthode TOBMI (Trans-Omic Block Missing Data Imputation) (297) essaye de pallier ce problème en utilisant tous les échantillons pour l'imputation, en ajoutant un poids supérieur aux plus proches voisins.

D'autres méthodes, telles que MI-MFA (Multiple Imputation in Multiple Factor Analysis) (298) propose une approche différente, en remplaçant aléatoirement les échantillons manquants par des échantillons d'un même phénotype n fois, résultant en plusieurs jeux de données complets différents, qui sont alors traités par MFA (Multiple Factor Analysis) (299). La méthode MFA est basée sur l'ACP, et cherche les similarités entre plusieurs matrices. Avec les nouvelles dimensions obtenues, tous les jeux de données créés sont concaténés en une configuration finale. Cette méthode s'affranchit de la recherche des plus proches voisins, mais peut utiliser des échantillons aberrants pour l'imputation. Des méthodes basées sur de l'apprentissage profond, telles que DeepMF (300), cherchent à reconstruire les matrices d'origine après les avoir compressées. Ils se servent de ces nouvelles matrices pour imputer les valeurs manquantes. Cependant, comme toutes méthodes d'apprentissage automatique, ils requièrent un certain nombre d'échantillons pour permettre un bon apprentissage. Enfin, un nombre limité de méthodes d'intégration multi-omiques autorisent une faible proportion d'échantillons manquants. C'est le cas de MOFA (279,288), qui arrive à décomposer les matrices initiales pour obtenir des facteurs malgré quelques valeurs absentes.

5.2.2. La malédiction de la dimension

Compte tenu du coût parfois élevé des techniques d'analyses à grande échelle, et/ou de la rareté de certains échantillons, il est fréquent d'avoir une cohorte d'échantillons de petite taille. Cependant, un nombre limité de données peut être insuffisant pour des analyses aussi complexes que des intégrations multi-omiques, au regard du grand nombre de caractéristiques. C'est ce qu'on appelle la malédiction de la dimension (curse of dimensionality), lorsque le nombre de variables est beaucoup plus grand que le nombre d'observations. Le nombre d'observations est alors insuffisant pour capturer la complexité des variables ou pour détecter des phénomènes rares. Le calcul de la puissance statistique permet d'obtenir une évaluation afin de déterminer si la taille du jeu de données est adéquate. La puissance statistique est la probabilité qu'un test statistique détecte un effet réel, s'il existe.

Récemment, Taraziba S. et al. ont développé l'outil MultiPower (301), permettant d'effectuer une estimation de la puissance statistique d'un jeu de données multi-omiques.

Afin de réduire le problème de dimensionnalité, deux solutions sont alors possibles. La première est de réduire les dimensions du jeu de données multi-omiques en réduisant le nombre de caractéristiques. À ce propos, plusieurs méthodes sont envisageables, de manière supervisée et non supervisée. En approche supervisée, l'utilisation de tests statistiques visant à sélectionner les caractéristiques les plus associées au(x) phénotype(s) d'intérêt(s) sont applicables, telles que les tests de Student/Wilcoxon ou Anova/Kruskal Wallis pour les données continues, ou Khi2/Fisher pour les données catégorielles. D'autres algorithmes plus complexes, tels que mRMR (Maximum Relevance and Minimum Redudancy) (302), visent à sélectionner les caractéristiques les plus pertinentes et les moins redondantes à l'aide d'approches d'apprentissage automatique. Dans la même idée, la méthode RFE (Recursive Feature Elimination) (303) consiste à utiliser différentes méthodes d'apprentissage automatique pour prédire le phénotype des échantillons, conserver les caractéristiques les plus prédictives, et recommencer ce processus de manière itérative.

Un classement de toutes les caractéristiques est retourné, permettant de sélectionner les plus intéressantes. De manière non supervisée, les algorithmes de réduction de dimension, tels que l'ACP, sont applicables afin de conserver les caractéristiques les plus représentatives des différents axes de variabilité dans les différents jeux de données. Enfin, il est également possible de réduire le nombre de caractéristiques, en ne gardant que celles qui ont le plus de sens biologiques, comme celles impliquées dans des processus d'intérêt, associées à d'autres omiques (ex. ARN et protéines), etc...

La deuxième possibilité est non pas de réduire le nombre de caractéristiques, mais d'augmenter le nombre d'échantillons. Compte tenu de la difficulté financière ou technique d'effectuer les mêmes analyses omiques sur de nouveaux échantillons, de nombreuses méthodes visant à créer des échantillons synthétiques ont été développées ces dernières années. Par exemple, SMOTE (Synthetic Minority Over-sampling TEchnique) (304) est une approche initialement développée pour réduire le déséquilibre entre le nombre d'échantillons dans deux classes différentes, notamment pour les analyses en apprentissage automatique. Elle se base sur les K-plus proches voisins pour créer des échantillons synthétiques proches d'échantillons réels.

Concrètement, en utilisant deux échantillons similaires comme point de départ, elle crée un troisième échantillon synthétique où les variables sont obtenues en prenant une moyenne pondérée aléatoirement (avec des poids variables entre les deux échantillons) des variables des deux échantillons réels (**figure 36**). De nombreuses variantes de cette méthode ont été développées afin de répondre à des problèmes spécifiques, comme ADASYN (305), mieux adaptée en cas d'hétérogénéité des échantillons d'une même classe. En effet, elle prend en compte la densité de la répartition des échantillons de la classe minoritaire afin de ne pas créer un échantillon synthétique improbable. Bien que la question des K-plus proches voisins puisse être problématique dans dans un jeu de données multi-vues (comme discuté précédemment), SMOTE a déjà été utilisé dans le cadre d'intégrations multi-omiques (306). Une alternative appelée KNNOR (K-Nearest Neighbor OveRsampling approach) (307) adopte une approche itérative basée sur le principe de SMOTE. Dans cette méthode, un échantillon synthétique temporaire est généré entre deux échantillons réels proches, puis un autre échantillon synthétique est créé en combinant le premier échantillon synthétique avec un troisième échantillon réel proche. Cette procédure est répétée n fois jusqu'à obtenir l'échantillon synthétique final (**figure 37**).

Cette approche permet d'éviter la création d'échantillons synthétiques à partir de seulement deux échantillons réels, mais plutôt à partir d'une multitude d'échantillons réels. Ainsi, davantage de variabilité est ajoutée aux données tout en restant dans l'espace des échantillons du même phénotype.

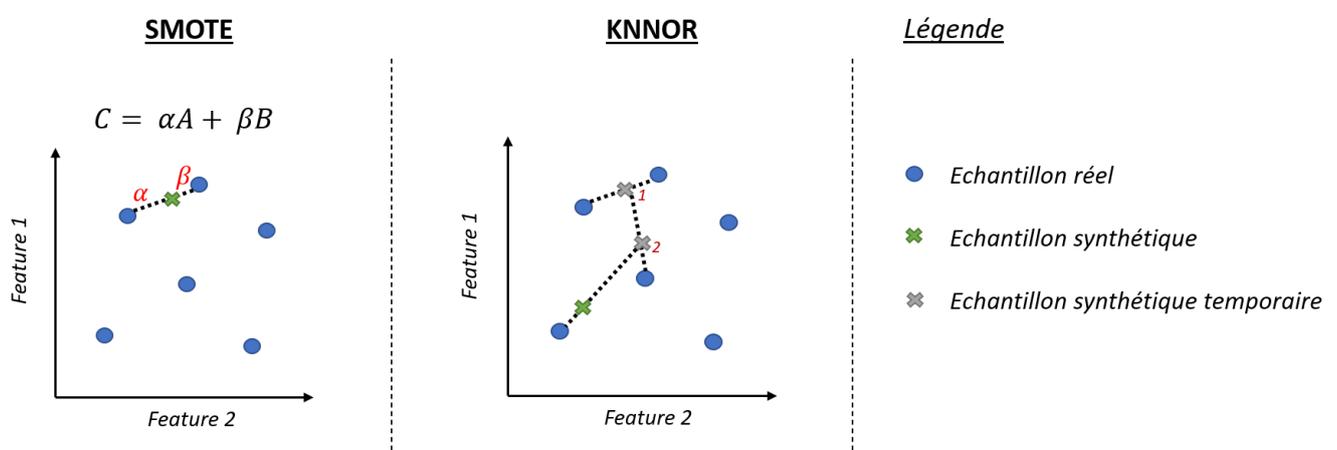


Figure 37. Représentation schématique du fonctionnement des méthodes SMOTE et KNNOR.

Enfin, des approches plus récentes utilisent des réseaux neuronaux afin d'apprendre la structure latente des données, et s'en servir pour créer des échantillons synthétiques (308). Ces derniers suivent donc la distribution apprise pendant l'entraînement et tendent à modéliser des échantillons qui pourraient hypothétiquement être réels. Ces approches ont le mérite de s'affranchir de l'approche des K-plus proches voisins et de créer des échantillons plus diversifiés, mais nécessitent un grand nombre de données d'entrée pour permettre un apprentissage suffisant.

5.2.3. L'hétérogénéité des données

Une autre difficulté inhérente aux intégrations multi-omiques est l'hétérogénéité des données. Par exemple, les données d'expression génique sont continues, tandis que les informations sur les mutations sont catégorielles. Leur modélisation mathématique sera donc totalement différente. De plus, des données continues ne suivent pas nécessairement les mêmes lois de probabilités : des expressions géniques ont tendance à suivre une loi binomiale négative, tandis que des données de méthylation suivent une distribution bimodale. De ce fait, le choix des méthodes d'intégration pour ces données est primordial (paragraphe 5.2.4).

Par ailleurs, la différence de taille des jeux de données peut entraîner une sur-représentation d'une omique, et biaiser le modèle d'intégration. Là où des données d'expression géniques peuvent contenir plus de 20.000 caractéristiques, des données de miRNAseq auront un nombre bien plus restreint, souvent inférieur à 1.000 caractéristiques. Des méthodes visant à réduire le nombre de caractéristiques peuvent être appliquées (comme discuté précédemment) afin d'équilibrer la taille des jeux de données omiques.

Enfin, l'association entre les caractéristiques de différentes omiques n'est pas toujours évidente. Par exemple, une relation entre un gène codant pour une protéine peut être assez claire, mais il est plus complexe d'interpréter la relation entre un site de méthylation situé dans un gène et l'expression d'un autre gène. Toutes ces questions doivent être prises en considération dans la mise en place de l'intégration multi-omiques et lors de l'interprétation des résultats obtenus.

5.2.4. Le choix de la méthode

Le choix de la méthode d'intégration multi-omiques peut avoir un impact significatif sur les résultats obtenus. Ce choix est étroitement lié à la question biologique sous-jacente.

L'exploration des réseaux biologiques peut bénéficier de l'utilisation de méthodes qui intègrent des connaissances préexistantes, telles que PARADIGM ou iOmicPass, tandis que l'identification de caractéristiques qui co-varient dans des conditions spécifiques peut être guidée par des approches de réduction de dimensions telles que MOFA, JIVE, MCIA ou DIABLO.

Parmi les méthodes appartenant à une même catégorie, certaines peuvent être préférées par leur fonctionnement. Par exemple, JIVE décompose le jeu de données initial en trois matrices : la variation conjointe entre les omiques, la variation propre à chaque omique et le bruit résiduel. Il permet donc de distinguer les variations communes aux différents jeux de données omiques, et celles portées par une omique particulière. En ce qui concerne MCIA, il projette les données dans un espace de dimensions réduit où chaque omique est traité de manière équivalente. Il analyse ensuite les co-variances entre chaque caractéristique, accordant un poids similaire aux caractéristiques partageant des tendances similaires. Ainsi, le jeu de données multi-omiques est transformé dans un espace commun où les dimensions représentent des axes de co-variances entre les caractéristiques. Ces axes peuvent représenter des conditions particulières des échantillons (âge, sexe, pathologie...), offrant ainsi la possibilité d'explorer en détail les expressions conditions spécifiques à ces situations.

MOFA décompose le jeu de données multi-omiques initial en deux matrices : celle des facteurs/individus (commune à chaque omique), et celle des caractéristiques/facteurs (spécifique à chaque omique). Chaque facteur est donc porteur à la fois d'un groupe d'échantillons partageant un fond génétique commun et d'un groupe de caractéristiques ayant un comportement similaire. Enfin, DIABLO diverge des trois méthodes précédentes en réalisant une factorisation supervisée, intégrant les méta-données des échantillons comme une matrice supplémentaire. Ainsi, l'objectif de DIABLO est d'expliquer les relations entre les caractéristiques associées au(x) phénotype(s) d'intérêt(s). Bien que reposant sur un principe général similaire, ces 4 méthodes présentent chacune des spécificités. MCIA et JIVE sont très proches, bien que JIVE distingue également les variations propres à chaque omique. À travers son analyse de facteurs, MOFA permet d'identifier des groupes de caractéristiques et/ou d'individus similaires, spécifiques à une ou plusieurs omiques. Il a également pour avantage de pouvoir intégrer à la fois des données continues et catégorielles. DIABLO, quant à lui, favorise les caractéristiques présentant des comportements propres aux méta-données des échantillons ([tableau 6](#)).

| Méthode | Intégration | Type de données | Gestion des valeurs manquantes | Variations propres à chaque omique |
|---------|----------------|--------------------------|--------------------------------|------------------------------------|
| JIVE | Non supervisée | Continue | Non | Oui |
| MCIA | Non supervisée | Continue | Non | Non |
| DIABLO | Non supervisée | Continue et catégorielle | Oui | Oui |
| MOFA | Supervisée | Continue | Non | Non |

Tableau 6. Comparaison des méthodes de réduction de dimension présentées ci-dessus.

En plus de celles mentionnées, de nombreuses autres méthodes sont disponibles, et plusieurs revues ont synthétisé et expliqué ces méthodes de manière concise pour faciliter leur sélection (280,296,309–311). La page GitHub "awesome-multi-omics" (312) répertorie également de nombreuses méthodes, ainsi que des revues et des exemples d'applications. Enfin, des outils comme Momix (313) permettent de tester simultanément un grand nombre de méthodes pour déterminer celles qui conviennent le mieux à un ensemble de données spécifique.

5.3. La complexité associée à l'interprétation des intégrations multi-omiques

Au-delà des difficultés techniques inhérentes aux intégrations multi-omiques, leur interprétation peut également être complexe. Les quatre méthodes de factorisation matricielles présentées ci-dessus ont été sélectionnées à titre d'exemple. La majorité de leurs applications se concentrent sur l'identification de set de caractéristiques d'intérêt, et/ou de processus biologiques dans lesquels ces caractéristiques sont impliquées. Dans cet objectif, Strefeler A. et al. ont utilisé MOFA pour étudier les dérégulations métabolomiques, lipidomiques et protéomiques dans le tissu cérébral de souris atteintes de la maladie d'Alzheimer et porteuses de mutations d'intérêt (314). Ils ont identifié 3 facteurs capturant des différences entre les souris sauvages et mutées, et/ou en mâles et femelles. Ils ont alors effectué un enrichissement fonctionnel des caractéristiques les plus impliquées dans ces facteurs, mettant en évidence des altérations spécifiques au sexe dans les métabolismes des lysophospholipides et des acides aminés.

| <i>Méthode</i> | Nombre de citations dans des études biologiques | Identification de caractéristiques et/ou de processus biologiques | Compréhension approfondie des mécanismes biologiques |
|-------------------|---|---|--|
| <i>JIVE</i> | 154 | 5 | 1 |
| <i>MClA</i> | 102 | 11 | 1 |
| <i>MOFA/MOFA+</i> | 413 | 42 | 8 |
| <i>DIABLO</i> | 205 | 75 | 9 |

Tableau 7. Application des méthodes multi-omiques présentées ci-dessus dans la compréhension de processus biologiques (juillet 2023).

Seules 14% de ces méthodes (19/133) ont permis une compréhension approfondie des mécanismes biologiques, en mettant en évidence des potentielles interactions entre des caractéristiques clé ([tableau 7](#)).

Par exemple, Sudhakar P. et al. ont exploité des données d'expressions spécifiques pour des types cellulaires particuliers (lymphocytes T CD4⁺ et monocytes CD14⁺) ainsi que des informations de génotypage (SNP) pour comprendre les mécanismes sous-jacents à la maladie de Crohn (315). L'utilisation de MOFA leur a permis d'isoler sept facteurs liés à au moins l'un des 13 caractères cliniques d'intérêt. Un clustering réalisé à partir des caractéristiques fortement associées à ces facteurs a permis de confirmer leur importance dans la pathologie, permettant bien de regrouper les échantillons par phénotype clinique. Les auteurs ont alors construit des réseaux de co-expression génétique pour chaque facteur. Cela leur a permis d'identifier des hubs, c'est-à-dire les gènes les plus interconnectés au sein de chaque réseau. Parmi ceux-ci se trouvaient des facteurs de transcription potentiellement impliqués dans les dérégulations génétiques propres à la maladie, ainsi que des cibles thérapeutiques déjà connues dans d'autres pathologies inflammatoires. Bhatt RR et al., quant à eux, travaillent sur les liens entre la psychologie de patients atteints de maladies inflammatoires chroniques de l'intestin et l'amélioration de leurs symptômes (316). En intégrant des informations relatives à la structure et à la connectivité cérébrale, ainsi que de nombreuses données cliniques et psychologiques, les auteurs ont identifié grâce à DIABLO des associations entre l'activation de zones

cérébrales liées à l'attention à la douleur et à l'amélioration des symptômes. En utilisant ces données, ils ont pu prédire l'amélioration du statut des femmes à 91% après 3 mois, et 83% après 12 mois. Ils en ont conclu que l'attention cérébrale portée à la douleur avait un impact significatif sur l'absence d'évolution chez les patients dont les symptômes n'évoluent pas.

La quantité d'informations générées par les méthodes d'intégration multi-omiques précoce influe sur la complexité de l'intégration qui en résulte. Cette complexité contribue en partie aux défis intrinsèques de ces approches, ce qui complique la caractérisation des processus biologiques à différentes échelles.

En conclusion, ce chapitre a abordé les avantages, mais également les difficultés inhérentes aux méthodes d'intégration multi-omiques. Bien que ces méthodes offrent des perspectives prometteuses pour mieux comprendre les processus biologiques, peu d'études s'engagent, à l'heure actuelle, dans une analyse complète et approfondie des phénomènes biologiques d'intérêt, nécessitant des approches spécifiques et un temps d'interprétation non négligeable.

Chapitre 6 : Mon projet de thèse

L'infection par le parasite *T. cruzi*, à l'origine de la maladie de Chagas, induit une réaction immunitaire inflammatoire ainsi qu'une production de ROS/RNS, contribuant toutes deux à son élimination. Toutefois, ces réponses ne suffisent pas à détruire le parasite, qui se réfugie dans différents organes, incluant le cœur. Chez 30% des individus, une cardiomyopathie se développe, caractérisée par une réponse inflammatoire exacerbée, et des dysfonctions mitochondriales. Des études antérieures ont mis en évidence une dérégulation génétique spécifique, potentiellement altérée par des mutations de l'ADN ou différents mécanismes épigénétiques (méthylation, ARN non codants). Cependant, le processus pathogénique demeure méconnu. Mon projet de thèse consiste à caractériser les processus biologiques impliqués dans les cardiomyopathies chagasiennes chroniques, et d'identifier les potentielles contributions génétiques et épigénétiques pouvant les affecter. Cela se traduit en plusieurs questions :

- Existe-t-il un profil génétique impliqué dans le développement des CCC ?
- Quels sont les mécanismes épigénétiques associés aux CCC et comment contribuent-ils à la perturbation observée dans l'expression génétique ?
- Comment tous ces éléments contribuent-ils conjointement à la pathogenèse des CCC ?

Afin de répondre à ces questions, j'ai contribué à l'exploration du profil génétique de patients atteints de CCC modérée et sévère, de patients asymptomatiques et de contrôles sains. Plusieurs études ayant déjà été réalisées sur les variants nucléaires, je me suis focalisée ici sur les variants mitochondriaux. Ce travail avait pour but d'évaluer l'importance de mutations mitochondriales dans le développement de cette pathologie.

J'ai également participé à l'analyse de données miRnomiques et transcriptomiques de patients CCC et de contrôles sains, afin d'étudier l'impact potentiel des miARN sur l'expression génique. De plus, j'ai mis en place une seconde approche épigénétique en essayant d'associer des données de méthylation avec des données transcriptomiques issues de séquençage à haut débit, permettant d'identifier non seulement des gènes codants mais également des ARNnc.

Cette étude avait pour objectif de caractériser les processus biologiques impactés par des différences de méthylation et d'expression génique, mais également d'estimer l'impact de la méthylation sur la dérégulation génétique associée aux CCC.

Pour finir, j'ai cherché à comprendre les liens entre tous ces éléments par le biais d'une intégration multi-omiques, incluant des données de génomique, transcriptomique, méthylomique et miRnomique. Cette approche globale visait à apporter une compréhension plus approfondie de la relation entre ces diverses informations et à explorer les mécanismes sous-jacents à la pathogenèse complexe des CCC.

RESULTATS

1. Le profil génétique des mitochondries est associé au développement des cardiomyopathies chagasiques chroniques

Article 1 : Mitochondrial DNA haplogroups and variants predispose to Chagas disease cardiomyopathy.

Frederic Gallardo*, **Pauline Brochet***, David Goudenège, João Paulo Silva Nunes, Pauline Andrieux, Barbara Maria Ianni, Amanda Farage Frade, Charles Mady, Ronaldo Honorato Barros Santos, Andreia Kuramoto, Samuel Steffen, Antonio Noedir Stolf, Pablo Pomerantzeff, Alfredo Inacio Fiorelli, Edimar Alcides Bocchi, Cristina Wide Pissetti, Bruno Saba, Fabrício C Dias, Marcelo Ferraz Sampaio, Fabio Antônio Gaiotto, José Antonio Marin-Neto, Abílio Fragata, Ricardo Costa Fernandes Zaniratto, Sergio Siqueira, Giselle De Lima Peixoto, Fernando Bacal, Paula Buck, Rafael Ribeiro Almeida, Hui Tzu Lin-Wang, André Schmidt, Mario Hiroyuki Hirata, Eduardo Antonio Donadi, Alexandre Costa Pereira, Virmondes Rodrigues Junior, Martino Martinelli, Michel Naslavsky, Jorge Kalil, Vincent Procaccio**, Edecio Cunha-Neto**,c, Christophe Chevillard**,c.

Soumis à Hearts.

Présentation de la publication :

Compte tenu de l'impact des mitochondries dans la maladie de Chagas, et notamment dans les CCC, notre équipe s'est focalisée sur l'étude du profil génétique de patients asymptomatiques, chagasiques modérés, chagasiques sévères, et de d'individus sains (contrôles). Avant tout, la quantité de mitochondries présente dans le tissu et le sang de patients a été quantifiée en évaluant l'expression de deux gènes marqueurs (ND1 et COX1) par PCR. Une réduction significative du nombre de mitochondries a été observée dans le tissu cardiaque de patients atteints de CCC comparé à des individus sains. En revanche, aucune variation n'a été constatée dans le sang d'individus asymptomatiques ou atteints de CCC modérés ou sévères. Cette variation du nombre de mitochondries est donc spécifique du tissu cardiaque.

Par la suite, l'ensemble des mutations présentes dans les mitochondries issues de tissu cardiaques d'individus sains (n = 6) et de patients atteints de CCC (n = 36) a été obtenu par séquençage direct (Hi-SNPseq) et annotées avec l'outil KerMit. Après reconstruction des séquences d'ADN contenant les mutations des différents individus, nous avons étudié les haplogroupes mitochondriaux des individus (lignées ancestrales, soit les différentes populations humaines) avec l'outil MITOMASTER. D'après cette analyse, les lignées génétiques africaines et européennes pourraient être plus fréquentes chez les patients atteints de CCC. Au regard de la faible taille de cette cohorte, une seconde analyse a été menée chez 38 patients asymptomatiques et 74 CCC (34 modérés et 36 sévères). Une troisième cohorte, contenant les mutations obtenues par Whole Genome Sequencing sur du sang de 289 individus sains brésiliens, a été ajoutée afin d'augmenter la puissance statistique de notre analyse. Ces nouvelles données ont confirmé l'association significative entre la lignée Africaine et le développement de la maladie de Chagas. De plus, une minorité d'individus européens est atteinte de cette pathologie. Toutefois, en se concentrant sur les différents phénotypes de patients chagasiques, il a été observé que la lignée africaine est associée aux formes asymptomatiques alors que l'Européenne est majoritaire chez les CCC, suggérant un effet protecteur ou délétère des mutations associées à ces haplogroupes. Une recherche plus avancée a montré des tendances d'association de différents macro-haplogroupes africains avec les différents stades de la maladie, suggérant un effet protecteur du macro-haplogroupe L0 contre les formes cardiaques, et délétère du macro-haplogroupe L3. Concernant l'haplogroupe européen, des tendances d'association du macro-haplogroupe H avec les CCC ont été notées.

Pour aller plus loin, nous nous sommes concentrés en détail sur les variants séquencés chez les individus chagasiques, et identifiés 712 variants significativement associés à une des formes de la maladie (asymptomatique, CCC modérée ou CCC sévère). 84 d'entre eux sont associés à un changement de la séquence protéique, tous annotés comme pathogéniques dans différentes bases de données, et 4 sont connus pour être impliqués dans des pathologies cardiaques. Ces mutations sont localisées dans 12 gènes mitochondriaux, notamment MT-ND5 (n = 21), MT-CYB (n = 15) et MT-ATP6 (n = 11).

L'ensemble de ces résultats tend à confirmer l'importance du profil génétique des mitochondries dans le développement de la maladie de Chagas, et notamment des cardiomyopathies. Toutefois, les mécanismes liant ces variants à la pathogénicité demeurent inconnus, et des études complémentaires sont nécessaires.

Contribution personnelle :

Dans cet article, dont je suis co-premier auteur, je me suis occupée des analyses statistiques à partir des fichiers VCF, contenant toutes les mutations par individu. Ces analyses incluent à la fois l'identification des haplogroupes des patients, ainsi que l'association de variants aux différentes formes de la maladie de Chagas. J'ai également contribué à la rédaction de l'article.

Mitochondrial DNA haplogroups and variants predispose to Chagas disease cardiomyopathy

Frederic Gallardo^{1,*}, Pauline Brochet^{1,*}, David Goudenège², João Paulo Silva Nunes^{1,3,4,5}, Pauline Andrieux¹, Barbara Maria Ianni³, Amanda Farage Frade^{3,4,5}, Charles Mady³, Ronaldo Honorato Barros Santos⁶, Andreia Kuramoto³, Samuel Steffen^{6,7}, Antonio Noedir Stolf⁷, Pablo Pomerantzeff⁸, Alfredo Inacio Fiorelli⁷, Edimar Alcides Bocchi⁷, Cristina Wide Pissetti⁹, Bruno Saba¹⁰, Fabrício C Dias¹¹, Marcelo Ferraz Sampaio¹⁰, Fabio Antônio Gaiotto^{6,7}, José Antonio Marin-Neto¹¹, Abílio Fragata¹⁰, Ricardo Costa Fernandes Zaniratto³, Sergio Siqueira¹², Giselle De Lima Peixoto¹², Fernando Bacal⁶, Paula Buck⁸, Rafael Ribeiro Almeida^{3,4,5}, Hui Tzu Lin-Wang¹⁰, André Schmidt¹¹, Mario Hiroyuki Hirata¹⁵, Eduardo Antonio Donadi¹¹, Alexandre Costa Pereira⁸, Virmondes Rodrigues Junior⁹, Martino Martinelli¹², Michel Naslavsky^{16,17}, Jorge Kalil^{3,4,5}, Vincent Procaccio², Edecio Cunha-Neto^{3,4,5,**,c} and Christophe Chevillard^{1,**,c}

¹ Institut National de la Santé Et de la Recherche Médicale (INSERM), Unité Mixte de Recherche (UMR) U1090, Aix Marseille Université, TAGC Theories and Approaches of Genomic Complexity, Institut MarMaRa, Marseille, France.

² Department of Genetics, University Hospital of Angers, 49933 Angers, France.

³ Laboratory of Immunology, Heart Institute Instituto do Coração (InCor), University of São Paulo, School of Medicine, São Paulo, Brazil.

⁴ Division of Clinical Immunology and Allergy, University of São Paulo, School of Medicine, São Paulo, Brazil.

⁵ Instituto Nacional de Ciência e Tecnologia, INCT, III- Institute for Investigation in Immunology, São Paulo, Brazil.

⁶ Division of Transplantation, Heart Institute Instituto do Coração (InCor), University of São Paulo, School of Medicine, São Paulo, Brazil.

⁷ Division of Surgery, Heart Institute Instituto do Coração (InCor), University of São Paulo, School of Medicine, São Paulo, Brazil.

⁸ Heart Institute (InCor), School of Medicine, University of São Paulo, São Paulo, São Paulo, Brazil.

⁹ Laboratory of Immunology, Universidade Federal Do Triângulo Mineiro (UFTM), Uberaba, Brazil.

¹⁰ Laboratório de Investigação Molecular em Cardiologia, Instituto de Cardiologia Dante Pazzanese (IDPC), São Paulo, Brazil.

¹¹ School of Medicine of Ribeirão Preto Faculdade de Medicina de Ribeirão Preto (FMRP), University of São Paulo, Ribeirão Preto, Brazil.

¹² Pacemaker Clinic, Heart Institute Instituto do Coração (InCor), School of Medicine, University of São Paulo, São Paulo, Brazil.

¹³ Heart Failure Unit, Heart Institute Instituto do Coração (InCor) School of Medicine, University of São Paulo, São Paulo, Brazil.

¹⁴ Department of Clinical and Toxicological Analyses, Faculty of Pharmaceutical Sciences, University of São Paulo (USP), São Paulo, Brazil.

¹⁵ Department of Clinical and Toxicological Analyses, Faculty of Pharmaceutical Sciences, University of São Paulo (USP), São Paulo, Brazil.

¹⁶ Human Genome and Stem Cell Research Center, Biosciences Institute, University of São Paulo, São Paulo, Brazil.

¹⁷ Hospital Israelita Albert Einstein, São Paulo, Brazil.

Name and complete address for correspondence: Edecio Cunha Neto, MD, PhD, Laboratório de Imunologia, Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina da Universidade de São Paulo. Av. Dr. Eneas de Carvalho Aguiar, 44 - Bloco II, 9º andar - 05403-000 São Paulo, SP, Brasil; Fax number: +55 11 2661-5953 / Telephone number: +55 11 2661-5914/5906 / E-mail address: edecunha@gmail.com; Christophe Chevillard, PhD, Theories and Approaches of Genomic Complexity (TAGC), INSERM, Aix Marseille University, UMR_1090, Parc Scientifique de Luminy, case 928, 163 avenue de Luminy, 13288 Marseille, France; Fax number: +33 491828701 / Telephone number: +33 491828736 / E-mail address: christophe.chevillard@univ-amu.fr

* Authors had an equal contribution.

** Authors had an equal contribution. 56
c Corresponding authors 57

Abstract: Cardiomyopathies are major causes of heart failure. Chagas disease (CD) is caused by the parasite *Trypanosoma cruzi*, and it is endemic in Central, South America. Thirty percent of the cases evolve into Chronic Chagas Cardiomyopathy (CCC) with worse prognosis as compared with other cardiomyopathies. In vivo bioenergetic analysis and ex vivo proteomic analysis of myocardial tissues highlighted worse mitochondrial dysfunction in CCC, and previous studies identified nuclear-encoded mitochondrial gene variants segregating with CCC. Here, we assessed the role of the mitochondrial genome through mtDNA copy number variations and mtDNA haplotyping and sequencing from heart or blood tissues of severe, moderate CCC and asymptomatic/indeterminate Chagas disease as well as healthy controls as an attempt to help decipher mitochondrial-intrinsic genetic involvement in Chagas disease development. We have found that mtDNA copy number was significantly lower in CCC than in heart tissue from healthy individuals, while blood mtDNA content was similar among asymptomatic Chagas disease, moderate and severe CCC patients. MtDNA haplogrouping study has indicated that African haplogroups were over represented in the Chagas subject groups in comparison with Brazilian healthy individuals. The European lineage is associated to protection against cardiomyopathy and the macro haplogroup H is associated with increased risk towards CCC. By mitochondria DNA sequencing, 84 mtDNA-encoded protein sequence pathogenic variants were associated with CCC. Among them, two variants were associated to left ventricular non-compaction and two to hypertrophic cardiomyopathy. The finding that mitochondrial protein-coding SNPs and mitochondrial haplogroups associate with risk of evolving to CCC is consistent with a key role of mitochondrial DNA in the development of Chronic Chagas disease Cardiomyopathy.

Keywords: chagas; cardiomyopathy; mitochondria; haplogroups; variants; copy-number

1. Introduction 85

Cardiomyopathies are a heterogeneous group of myocardial diseases associated with mechanical or electrical dysfunction that usually exhibit inappropriate ventricular hypertrophy or dilation. Dilated cardiomyopathy is one of the most common causes of heart failure (1) and the most common indication for heart transplantation worldwide, with an estimated prevalence of 40 cases per 100 000 individuals and an annual incidence of 7 cases per 100 000 individuals. Dilated cardiomyopathy occurs after several decades of life (2, 3). Moreover, adult males seem to be more frequently affected than adult females (2, 3). Approximately 20 to 35%, or potentially up to 50% of patient cases, could be attributed to the genetic basis (4). More than 100 genes have been associated to dilated cardiomyopathies (5-8). Myocardial energy deprivation is a main feature of cardiomyopathies and heart failure, and multiple evidences indicate this is caused by mitochondrial dysfunction. Because the heart is an organ requiring a high energy demand, regulation of mitochondrial metabolism plays an important role in the pathogenesis of dilated cardiomyopathies. During the first stages of dilated cardiomyopathies, increased mitochondrial proliferation acts as a compensating mechanism for maintaining energy supply (9). However, the number of mitochondria declines during dilated cardiomyopathy progression leading to a reduction in ATP, decreased contractility and increased reactive oxidative species (ROS) all of which resulting in diastolic dysfunction and heart failure (10). In many cases

mitochondrial cardiomyopathies have an underlying genetic component resulting in mitochondrial respiratory chain deficiencies (11-13), including fatty acid oxidation (11) or cardiolipin synthesis defects (14-16) and alterations of mitochondrial dynamics (11). Mitochondria are under dual genome control, where a small fraction of mitochondrial proteins are encoded by mitochondrial DNA (mtDNA), while more than 99% of them are encoded by nuclear DNA (nDNA). Mutations in over 250 mitochondria-related nDNA or mtDNA genes have been associated with mitochondrial disease, resulting in mitochondrial dysfunction leading to insufficient energy production required to meet the needs for organs with high energy requirements (17).

Cardiac involvement occurs in mitochondrial diseases, with cardiomyopathies being one of the most frequent cardiac manifestations found in these disorders occurring in 20–40% of patients. Mitochondrial cardiomyopathies can vary in severity from asymptomatic status to severe manifestations including heart failure, arrhythmias, and sudden cardiac death (11). MtDNA polymorphisms that have occurred over evolutionary time have permitted a distribution of mtDNA variants into mitochondrial haplogroups inherited from the maternal lineage, and these haplotypes may influence mitochondrial function (18).

Chagas disease (CD) is caused by the parasite *Trypanosoma cruzi*, and it is endemic in Central, South America, and Mexico. The clinical course of the disease comprises an acute phase, mostly asymptomatic, and a chronic phase, where 60% of the patients remain asymptomatic. Forty percent of the cases evolve into gastrointestinal disease and chronic cardiomyopathy with varying degrees of severity including refractory heart failure (19). Antiparasitic treatment with benznidazole or nifurtimox is most effective in the acute phase and prevents progression to chronic phase when given to children. However, these treatments had multiple adverse effects (20). Heart failure due to Chronic Chagas Cardiomyopathy (CCC) may have a worse prognosis with 50% shorter survival when compared to other cardiomyopathies of different etiologies such as ischemic cardiomyopathy and idiopathic dilated cardiomyopathy (21, 22). The major histopathological feature attending dilated cardiomyopathy in CCC is the presence of a diffuse myocarditis, with intense cardiomyocyte damage and hypertrophy, and significant fibrosis, in the presence of very scarce *T. cruzi* forms. The myocardial inflammatory infiltrate is thought to play a major role in disease development and progression (23, 24). The myocardial cytokine production profile suggests an IFN γ /TNF α Th1 type response, with interferon γ -induced chemokines (25-28). Moreover, our group has previously demonstrated that CCC myocardium presents a unique gene expression profile, distinct from other noninflammatory dilated cardiomyopathies, with a large number of interferon-gamma inducible genes among differentially expressed RNA (29). The proteomic analysis of myocardial tissues revealed that CCC display a specific global protein expression profile. Pathway analysis of proteins differentially expressed in CCC highlighted mitochondrial dysfunction, cardiac hypertrophy, fibrosis, mitochondrial energy metabolism, fatty acid metabolism, the involvement of Creatine Kinase System and ATP Synthase Complex, variations of the mitochondrial membrane potential (mitochondrial $\Delta\Psi_m$) (30-32). IFN- γ and TNF- α treatment of the human cardiomyocyte cell line AC16 induces a dose-dependent reduction of mitochondrial $\Delta\Psi_m$ and mtDNA copy number, suggesting a possible inflammatory origin for mitochondrial dysfunction in CCC (31). IFN- γ and TNF- α treatment of AC-16 cardiomyocytes' mitochondria also induce an increased nitrooxidative stress profile. It induces impairment of mitochondrial $\Delta\Psi_m$ (33). This finding is especially crucial since the mitochondrial $\Delta\Psi_m$ is the OXPHOS

proton motive force that drives ATP production through ATP synthase is an essential mechanism for contraction and survival of cardiac cells (34).

About 30% of Chagas disease patients develop CCC, suggesting the participation of modifier genes in differential disease progression. This was reinforced by the discovery of familial aggregation of cases of CCC in endemic area settings (35). Based on whole exome sequencing on multiple familial cases of Chagas disease, we have previously identified heterozygous, pathogenic variants linked to CCC in all tested families on 22 distinct nuclear-encoded genes, from which 20 were mitochondrial or inflammation-related - most of the latter involved in proinflammatory cytokine production (36). So, mitochondrial dysfunction and inflammation may be genetically determined in CCC, driven by rare genetic variants which increase mitochondrial susceptibility to IFN- γ -induced damage in the myocardium. Mutations in the mitochondrial genome accumulate over time in organs suffering mitochondrial dysfunction, resulting in decreased energy production and reactive oxygen species (ROS) overproduction. Defective copies of mtDNA might accumulate over time, raising the overall heteroplasmy level of deleterious mutations. Heteroplasmy is the concurrence of both normal (wt) and abnormal (mutant) mtDNA within one cell. In general, the level of mitochondrial dysfunction is commensurate with the mutational load of mtDNA within the tissue or cell.

In this paper, we assessed the role of the mitochondrial genome through mtDNA copy number variations and mtDNA haplotyping and sequencing from heart or blood tissues of severe, moderate CCC and asymptomatic/indeterminate Chagas disease as well as healthy controls as an attempt to help decipher mitochondrial-intrinsic genetic involvement in Chagas disease development.

2. Materials and Methods

2.1. Patients and Tissue Collection

Human left ventricular free wall heart tissue samples were obtained from patients with end-stage heart failure CCC at the time of heart transplantation (n = 34). These CCC patients underwent serological diagnosis of *T. cruzi* infection and standard electrocardiography and echocardiography, and tissues were subject to histopathological assessment. Moderate CCC patients are characterized by an ejection fraction ≥ 0.4 whereas severe CCC patients are characterized by an ejection fraction < 0.4 . Biopsies from controls (n = 6) were obtained from healthy hearts of organ donors having no suitable recipient. The protocol was approved by the institutional review boards of the University of São Paulo School of Medicine (n°009/2011, January 21th 2011) and INSERM (French National Institute of Health and Medical Research) (FWA00005831, March 16th 2011). Written informed consent was obtained from all patients. All experimental methods comply with the Helsinki Declaration. The protocol was approved by the institutional review boards of the University of São Paulo School of Medicine (n°009/2011, January 21th 2011) and INSERM (French National Institute of Health and Medical Research) (FWA00005831, March 16th 2011). Written informed consent was obtained from all patients. All experimental methods comply with the Helsinki Declaration.

2.2. Heart tissue and blood DNA isolation

Heart tissue samples were crushed twice with lysis buffer. After proteinase K treatment, DNA was extracted with QIAamp DNA Mini Kit (Qiagen, Courtaboeuf, France) according to the manufacturer's recommendations. For each subject, 5 to 15 ml of blood were collected into EDTA tube. Genomic DNA was isolated with FlexiGene DNA kit (Qiagen, Courtaboeuf, France) according to the manufacturer's recommendations.

2.3. Mitochondrial DNA copy number quantification 214

The Quantitative PCR (qPCR) was conducted using SYBR Green/ROX qPCR Master Mix (2x) (Thermo Scientific), 1 ng of template DNA and primers for mitochondrial-encoded genes NADH: ubiquinone oxidoreductase subunit 1 (MT-ND1), cytochrome c oxidase subunit 1 (MT-COXI). Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was used as the endogenous control. The ratio of mtDNA/nDNA was calculated using the formula $2^{2-\Delta\Delta CT}$ (37). The primer sequences were: MT-ND1 forward: 5' ATACCCATGGCCAACCTCCT 3', MT-ND1 reverse: 5' GGGCCTTTGCGTAGTTGTAT 3'; MT-COXI forward: 5' TCCACTATGTCCTATCAATA 3', MT-COXI reverse: 5'GGTGTAGCCTGA-GAATAG 3'; GAPDH forward: 5' CCCTGTCCAGTTAATTTTC 3', GAPDH reverse: 5' CACCCTTTAGGGAGAAAA 3'. 215
216
217
218
219
220
221
222
223
224
225

2.4. Mitochondrial haplogroup determination 226

Mitochondrial haplogroups were determined according to MITO-MASTER tool (<https://www.mitomap.org/foswiki/bin/view/MITO-MASTER/WebHome>) which used HaploGrep2 with Phylotree 17 for haplogroup determination. Mitochondrial DNA genotyping was assessed using four different methods: 227
228
229
230
231

PCR amplicon sequencing: The genomic DNA was amplified by PCR. The primer sequences were: HmtL15997: 5' CACCATTAGCACCCAAAGCT 3', HmtH112: 5' ACAGATACTGCGACATAGGG 3'. PCR amplifications were carried out with GoTaq polymerase (Promega, Charbonnières-les-Bains, France) and 1 μ M of each primer. 50 μ l reactions were carried out on Eppendorf thermocycler. The PCR products were loaded on agarose gel and purified with QIAEXII gel extraction kit (Qiagen) before Sanger sequencing. Sanger sequencing of the PCR amplicon was done with the same primers. 232
233
234
235
236
237
238
239

Direct sequencing: Hi-SNPseq, provided by CD Genomics, combines multiplex PCR and high-throughput sequencing (CD Genomics, Shirley, USA). Multiplex PCR amplification was done with 110 site-specific primers in two single tubes. These two reactions comprise the entire library covering the whole mtDNA sequence. Then, the amplicon pooled samples were purified to remove polymerase, primers, and short by-products. Different samples were distinguished by different barcoded primers. After mixing the samples, high-throughput sequencing was performed (pair ends 150). The sequences were aligned on the reference rCRS genome. 240
241
242
243
244
245
246
247
248

Whole Exome sequencing: Library preparation (Agilent SureSelect Human All Exon V6) and sequencing steps were commissioned to Genewiz (Leipzig, Germany). Raw BCL files generated by the sequencer were converted to fastq files for each sample using bcl2fastq v.2.19. Sequence reads were trimmed to remove possible adapter sequences and nucleotides with poor quality using Trimmomatic v.0.38. The sequences were aligned on the reference rCRS genome. 249
250
251
252
253
254
255

Whole genome sequencing: For the Brazilian reference population (38), high-coverage WGS data were aligned on the human reference hg38 genome. For each sample, variant calling was done using the Genome Analysis Toolkit 4 (GATK4). Based on the mitochondria variants a mitochondria specific bam file. 256
257
258
259
260

2.5. Variant annotations 261

Sequencing data analysis was performed using a homemade nextflow pipeline that integrates all the classical NGS steps for molecular diagnosis. Mapping and quality control was done using BWA (39) and FastQC (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) respectively. Base recalibration, base quality score recalibration and duplicate marking 262
263
264
265
266

processed with elPrep (40). Variant calling combines the results of GATK Mu- 267
tuct2 (41) in mitochondrial mode and Strelka2 (42). Variant annotation and 268
prioritization were performed using an in-house tool KerMit that integrates 269
various mitochondrial dedicated tools and databases like MITOMAP, 270
HmtDB, ClinVar, MitoTip and HaploGrep for haplogroup prediction. Mito- 271
chondrial deletions were searched using eKLIPse (43). 272

2.6. Statistical analysis 273

All the statistical tests have been made using R 3.6.3. A first comparison 274
on the presence/absence of genetic variants between two groups was carried 275
out using the chi2 and fisher tests. A second analysis, based on allelic fre- 276
quency, was done with a Kruskal Wallis test between all groups, and then 277
with a Dunn test comparing phenotypes two by two using the rstatix R pack- 278
age. All the plots have been done with the ggplot2 R package. 279

Differences between groups are evaluated by statistical tests (Chi2 test 280
with Yate's correction, Fisher's exact test, Kruskal Wallis test) using GraphPad 281
Prism 9.3.0. Asterisk and the sharp sign (*) indicate *p-value* <0.05, (***p-value* 282
<0.001, (***) *p-value* <0.0001. 283

3. Results 284

3.1. MtDNA copy number is reduced in heart of Chagas patients 285

The quantification of mtDNA copy number was evaluated on heart tissue 286
samples from end stage patient heart tissues (n=33) and on healthy hearts of 287
organ donors having no suitable recipient (n=6) (Supplementary Table S1). 288
MtDNA copy number using (MT-ND1) was significantly lower in CCC ((MT- 289
ND1 69%; *p*= 0.0034) than control samples (Figure 1A) indicating a reduction 290
of mtDNA copy number. Mitochondrial DNA (MT-COX1) gave a similar re- 291
sult (37%; *p*= 0.0019) (Figure 1B). 292

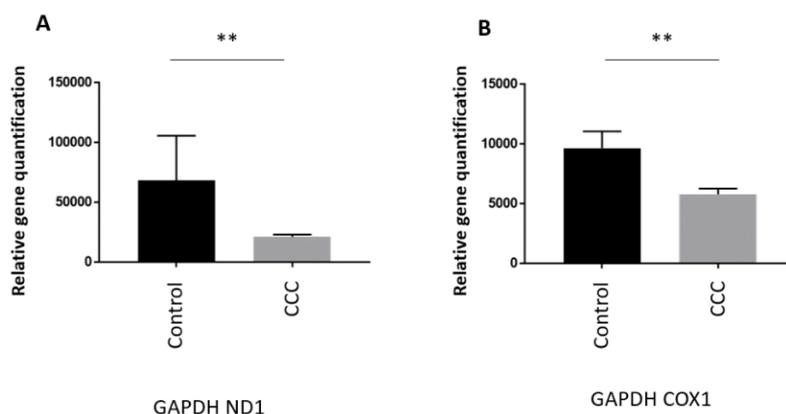


Figure 1. Mitochondrial DNA quantification in heart tissues. We performed qPCR with 294
primers for mitochondrial encoded genes NADH: ubiquinone oxidoreductase subunit 295
1 (MT-ND1), cytochrome c oxidase subunit 1 (MT-COX1) and endogenous nuclear-en- 296
coded control gene Glyceraldehyde-3-phosphate dehydrogenase (GAPDH). The rela- 297
tive quantification of mitochondrial DNA was calculated by determining the $\Delta C_t =$ 298
 $C_t(\text{endogenous nuclear-encoded control gene}) - C_t(\text{mitochondrial encoded genes})$ and 299
then using the formula $2^{x\Delta C_t}$. 300

The mtDNA copy number was also evaluated in DNA extracted from 301
blood of asymptomatic subjects (n=209), moderate CCC patients (n=235) and 302
severe CCC patients (n=168) (Supplementary Table S2). The age of the indi- 303
viduals which ranges from 40 to 60 years old, was similar among the three 304

groups (Figure 2A). As shown on Figure 2B and 2C, the mtDNA content (and therefore mitochondrial mass) was similar between the three groups. So, the decrease of the mitochondrial mass detected on heart biopsies seems to be tissue specific.

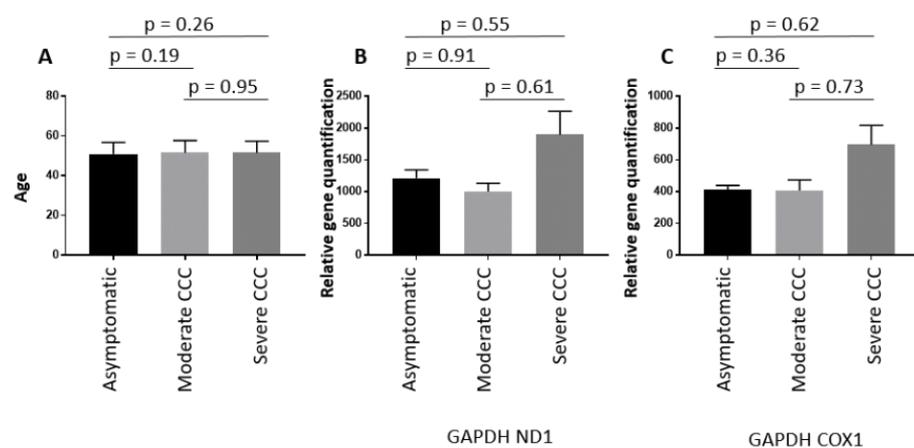


Figure 2. Mitochondrial DNA quantification in blood DNA samples.

This quantification was done on blood DNAs from patients between 40 and 60 years old. These patients were either asymptomatic for Chagas disease, or developed chronic cardiomyopathy stratified as moderate or severe based on left ventricular ejection fraction values [moderate CCC ($0.40 \leq EF \leq 1$) and severe CCC ($0 \leq EF < 0.40$)].

Human mtDNA haplogroups and their distributions have been extensively investigated across different nations and geographical regions, predominantly to define population origins and genetic structure. The rate at which mtDNA mutates is known as the mitochondrial molecular clock. The entire mtDNA was sequenced from 26 end stage heart tissue samples and 6 healthy hearts of organ donors. Table 1 describes the mitochondrial haplogroup lineage distribution between both groups.

Table 1. Haplogroup distribution in heart tissue samples.

| Haplogroup lineage | End stage heart tissue samples (CCC) | Healthy hearts of organ donors (control) |
|-----------------------|--------------------------------------|--|
| African | 19 (73%) | 4 (67%) |
| European | 4 (15%) | 0 (0%) |
| Native American/Asian | 2 (8%) | 1 (17%) |
| Eurasian | 1 (4%) | 1 (17%) |

This distribution suggests that African and European lineages may be enriched in end stage CCC patients compared to healthy controls. However, due to the limited size of our study group, the difference was not statistically significant.

In order to investigate this haplogroup distribution, a larger collection of blood DNA was then investigated. It included CCC patients (n=321), asymptomatic subjects (n=38) and we had access to a census-based sample of elderly individuals from São Paulo (Brazil) (Supplementary Table S3). This latest reference population includes 1171 healthy individuals (38). For each of them, the WGS was available and mitochondrial haplogrouping was assessed in a group of 289 individuals selected randomly from this cohort. The three main lineages were African, European and Native American/Asian (Figure 3). Chagas individuals (asymptomatic or with Chronic Chagas Cardiomyopathy)

were compared to healthy controls, percentage of African lineage in healthy individuals reached 27% whereas this percentage was 55% in the Chagas patient groups. The Chi-square test with Yates' correction was significant ($p < 0.0001$) with a relative risk of 2.004. This may imply that Chagas patients have a higher African mother lineage than healthy elderly individuals from the state of São Paulo, Brazil.

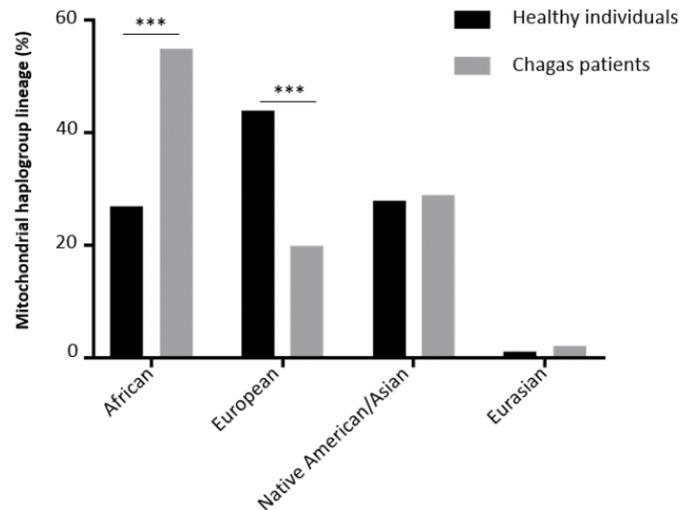
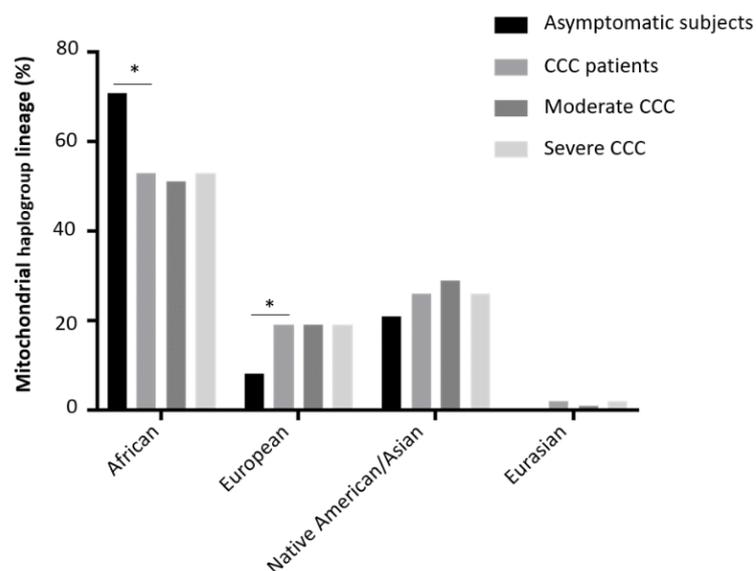


Figure 3. Mitochondrial haplogroup distribution between the healthy Brazilian reference population and the Chagas cohort.

The percentage of European lineages in healthy individuals reached 44% whereas this percentage was 18% in the Chagas subject groups. The Chi-square test with Yates' correction gave a significant p value ($p < 0.0001$) for a relative risk of 0.5292 (Figure 3).

We then stratified the Chagas subject group, in the asymptomatic/indeterminate subjects, the moderate CCC ($0.40 \leq EF \leq 1$) and severe CCC ($0 < EF < 0.40$) patients. The percentage of each lineage in each subgroup is shown on Figure 4. The African lineage was more represented in the asymptomatic sub-group than the CCC patient sub-group (71% vs 53%) ($p = 0.04$). Similarly, the European lineage was overrepresented in the CCC patient sub-group (7.9% vs 18.7%) but this difference was not significant ($p = 0.12$).



343

344

345

346

347

348

349

350

351

352

353

354

355

356

Figure 4. Mitochondrial haplogroup distribution among the Chagas cohort.

358

Focusing on African sub-haplogroups, the distribution was compared between ASY subjects, CCC patients, moderate CCC patients and severe CCC patients (Figure 5). This African sub-haplogroup distribution was not significantly different between the study groups (Supplementary Table S4). However, the percentage of L0 sub-haplogroup in ASY individuals reached 11% whereas this percentage was 4% in the CCC patient group. The percentage of L3 sub-haplogroup in ASY individuals reached 22% whereas this percentage was 33% in the CCC patient group. It may suggest that the L0 sub-haplogroup has a protective effect whereas the L3 sub-haplogroup may act as a disease mark.

359
360
361
362
363
364
365
366
367
368

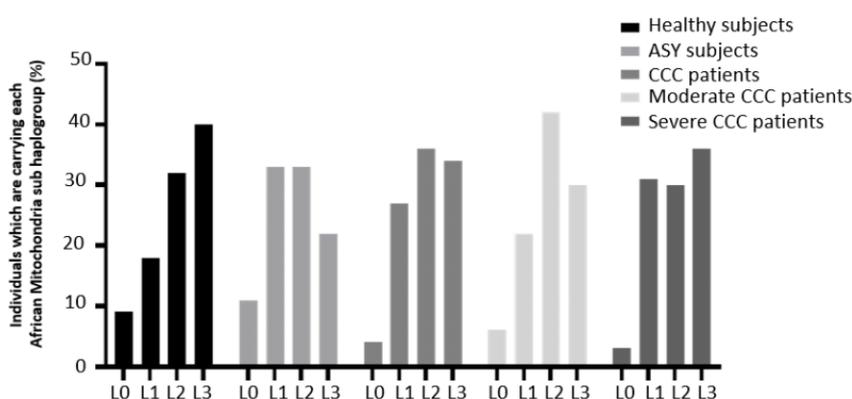


Figure 5. African mitochondrial sub-haplogroup distribution in the Chagas cohort.

369

In the same way, the European sub-haplogroup distribution was similar between the CCC study groups (Figure 6). However, when we compared the frequency of the European H sub-haplogroup, we detected some trends of association Table 2.

370
371
372
373

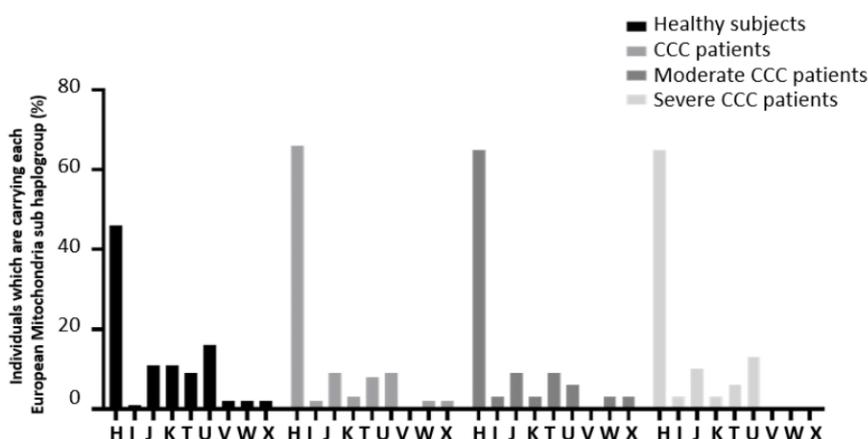


Figure 6. European mitochondrial sub-haplogroup distribution in the Chagas cohort.

374

375

Table 2. European haplogroup H association study.

376

| Groups | <i>p</i> Value | Relative Risk |
|--|----------------|---------------|
| Healthy individuals versus chagas patients | 0.01 | 0.76 |
| Healthy individuals versus moderate chagas cardiomyopathy patients | 0.06 | 0.85 |
| Healthy individuals versus severe chagas cardiomyopathy patients | 0.07 | 0.86 |

Whole mtDNA sequencing was carried out on 112 samples (Supplementary Table S5). It included 38 asymptomatic subjects and 74 CCC patients. Among the CCC patients, 34 patients suffered from moderate cardiomyopathy whereas 36 patients had a severe cardiomyopathy. For 4 CCC patients, the ejection fraction value is not available.”

Sequences obtained from on 112 samples were aligned according to the Cambridge Reference Sequence (rCRS). 12429 variants were detected. Three analyses were done. First of all, we compared the number of carriers of each variant between each group (ASY vs CCC; ASY vs Moderate CCC; ASY vs Severe CCC and Moderate CCC vs Severe CCC) (Supplementary Table S6). We also compared the heteroplasmy levels between the three groups (Kruskal Wallis between ASY, moderate CCC and severe CCC). Finally, we compared the heteroplasmy levels between the groups (two by two: ASY vs CCC; ASY vs Moderate CCC; ASY vs Severe CCC and Moderate CCC vs Severe CCC) (Supplementary Table S7). All these tests were corrected for multiple testing. 712 variants showed significant results in at least one test (adjusted p value \leq 0.05).

We annotated the associated variants, based on available information in VEP ClinVar, MitoMap and MitImpact databases (Supplementary Table S8 and S9). These variants are either located in intergenic region (n=220; 30.8%) or in coding region (n=493; 69.2%). Among these variants, only 84 variants change the protein sequence (Table 3). These protein sequence variants are enriched in a limited number of genes (MT-ND5 (n=21), MT-CYB (n=15), MT-ATP6 (n=11), MT-CO3 (n=8), MT-CO1 (n=7), MT-ND4 (n=6), MT-ATP8 (n=5), MT-ND6 (n=4) MT-ND4L (n=2), MT-ND1 (n=2), MT-CO2 (n=2), MT-ND2 (n=1)). The pathogenicity of each variant was defined with MitImpact which aggregate 24 databases (Clinvar, PolyPhen2, SIFT, FatHmW, FatHm, PROVEAN, Mutatiossessor, EFIN_SP, EFIN_HD, CADD, VEST, PANTHER, PhDSNP, SP, MutationTaster, SNPDryad, Condel, COVEC_WMVB, MtoolBox, APOGEE, DEOGEN2, PolyPhen2 transf, SIFT transf, Mutatiossessor transf) (44). All the 84 protein modifier variants were classified as pathogenic in at least one database. 16 protein modifier variants were classified as pathogenic in more than 10 databases.

377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410

Table 3. Annotation of the 84 protein modifier variants.

| Pos | Ref | Alt | Csq | Gene symbol | dbSNP ID | Uniprot name | Uniprot ID | AA pos | AA ref | AA alt | Described as pathogenic in the following databases. |
|------|-----|-----|-----|-------------|--------------|--------------|------------|--------|--------|--------|---|
| 3565 | A | C | RNA | MT-ND1 | | NU1M | P03886 | 87 | T | P | PolyPhen2, SIFT, PROVEAN, CADD, PhDSNP, SP, COVEC_WMV, MtoolBox |
| 3579 | A | C | RNA | MT-ND1 | | NU1M | P03886 | 91 | M | I | CADD, PhDSNP, MutationTaster, SNPDRyad, Condel |
| 5262 | G | A | RNA | MT-ND2 | | NU2M | P03891 | 265 | A | T | CADD, PANTHER, Condel, |
| 6571 | C | T | RNA | MT-COI | | COX1 | P00395 | 223 | A | V | PROVEAN, EFIN_SP, CADD, PhDSNP, MutationTaster, Condel, PolyPhen2 transf |
| 6873 | C | T | RNA | MT-COI | | COX1 | P00395 | 324 | L | F | PolyPhen2, SIFT, Mutatiossessor, EFIN_SP, EFIN_HD, CADD, PANTHER, PhDSNP, MutationTaster, COVEC_WMV, MtoolBox, Mutatiossessor transf |
| 6899 | G | C | RNA | MT-COI | rs1556423194 | COX1 | P00395 | 332 | M | I | Condel, PolyPhen2 transf, SIFT transf |
| 6909 | G | T | RNA | MT-COI | | COX1 | P00395 | 336 | A | S | SIFT, CADD, PhDSNP |
| 6915 | G | A | RNA | MT-COI | rs1603220687 | COX1 | P00395 | 338 | V | M | Condel, PolyPhen2 transf, SIFT transf |
| 7008 | G | T | RNA | MT-COI | | COX1 | P00395 | 369 | D | Y | PolyPhen2, SIFT, FatHmm, PROVEAN, Mutatiossessor, EFIN_SP, EFIN_HD, CADD, PANTHER, PhDSNP, SP, MutationTaster, SNPDRyad, COVEC_WMV, MtoolBox, APOGEE, Mutatiossessor transf |
| 7072 | T | C | RNA | MT-COI | rs1603220760 | COX1 | P00395 | 390 | M | T | PolyPhen2, SIFT, EFIN_HD, CADD, PhDSNP, SP, MutationTaster, COVEC_WMV, MtoolBox |
| 7868 | C | T | RNA | MT-CO2 | rs1556423357 | COX2 | P00403 | 95 | L | F | Condel, PolyPhen2 transf |
| 8021 | A | G | RNA | MT-CO2 | rs1603221261 | COX2 | P00403 | 146 | I | V | PROVEAN, CADD, PANTHER, PhDSNP, Condel, PolyPhen2 transf |
| 8401 | A | C | INT | MT-ATP8 | | ATP8 | P03928 | 12 | M | I | Condel, PolyPhen2 transf |
| 8405 | A | G | INT | MT-ATP8 | rs1603221462 | ATP8 | P03928 | 14 | T | A | Condel, SIFT transf |
| 8527 | A | G | RNA | MT-ATP6 | rs878853003 | ATP6 | P00846 | 1 | M | V | Condel, PolyPhen2 transf |
| 8558 | C | G | RNA | MT-ATP8 | | ATP8 | P03928 | 65 | P | A | PolyPhen2, FatHmm, PROVEAN, CADD, SP, COVEC_WMV, MtoolBox |
| 8558 | C | G | RNA | MT-ATP8 | | ATP6 | P00846 | 11 | A | G | CADD, PANTHER, PhDSNP |
| 8566 | A | G | RNA | MT-ATP6 | rs3020563 | ATP6 | P00846 | 14 | I | V | SP, Condel |
| 8568 | C | T | RNA | MT-ATP8 | rs1603221589 | ATP8 | P03928 | 68 | S | F | PolyPhen2, FatHmm, EFIN_HD, CADD, SP, MtoolBox |
| 8618 | T | C | RNA | MT-ATP6 | rs28358885 | ATP6 | P00846 | 31 | I | T | Condel |
| 8632 | T | C | RNA | MT-ATP6 | rs1603221654 | ATP6 | P00846 | 36 | Y | H | PANTHER, Condel, PolyPhen2 transf |
| 8677 | A | C | RNA | MT-ATP6 | | ATP6 | P00846 | 51 | K | Q | Condel, PolyPhen2 transf |
| 8714 | C | T | RNA | MT-ATP6 | rs1603221724 | ATP6 | P00846 | 63 | T | I | FatHmm, PANTHER, Condel |
| 8795 | A | C | RNA | MT-ATP6 | | ATP6 | P00846 | 90 | H | P | PolyPhen2, SIFT, PROVEAN, CADD, PANTHER, PhDSNP, SP, SNPDRyad, COVEC_WMV, MtoolBox |
| 8854 | G | A | RNA | MT-ATP6 | rs386829055 | ATP6 | P00846 | 110 | A | T | SIFT, CADD, PANTHER, PhDSNP, Condel |
| 8860 | A | G | RNA | MT-ATP6 | rs2001031 | ATP6 | P00846 | 112 | T | A | PROVEAN, PhDSNP, SP, Condel, PolyPhen2 transf |
| 8870 | T | C | RNA | MT-ATP6 | rs1556423560 | ATP6 | P00846 | 115 | M | T | Condel, PolyPhen2 transf |
| 8875 | T | C | RNA | MT-ATP6 | rs201123510 | ATP6 | P00846 | 117 | F | L | Condel, PolyPhen2 transf |
| 9261 | A | C | RNA | MT-CO3 | | COX3 | P00414 | 19 | T | P | Condel, PolyPhen2 transf |
| 9325 | T | C | RNA | MT-CO3 | rs879000531 | COX3 | P00414 | 40 | M | T | PolyPhen2, PROVEAN, EFIN_SP, CADD, PhDSNP, SP, SNPDRyad, MtoolBox |
| 9717 | C | G | RNA | MT-CO3 | | COX3 | P00414 | 171 | L | V | Condel, PolyPhen2 transf |
| 9861 | T | C | RNA | MT-CO3 | rs878853060 | COX3 | P00414 | 219 | F | L | Condel, PolyPhen2 transf, SIFT transf |
| 9877 | T | A | RNA | MT-CO3 | | COX3 | P00414 | 224 | M | K | MutationTaster, Condel |
| 9880 | T | A | RNA | MT-CO3 | | COX3 | P00414 | 225 | F | Y | Condel, PolyPhen2 transf |
| 9896 | A | C | RNA | MT-CO3 | | COX3 | P00414 | 230 | K | N | Condel, PolyPhen2 transf |

| 9924 | T | G | RNA | MT-CO3 | COX3 | P00414 | 240 | W | G | PolyPhen2, PROVEAN, Mutatiosessor, EFIN_HD, CADD, VEST, PhDSNP, SP, SNPDryad, COVEC, WMV, MtoolBox |
|-------|---|---|-----|---------|-------|--------|-----|---|---|--|
| 10677 | G | A | RNA | MT-ND4L | NU4LM | P03901 | 70 | E | K | FatHmm, PROVEAN, Mutatiosessor, EFIN_SP, CADD, PANTHER, PhDSNP, SP, MutationTaster, SNPDryad, Condel, MtoolBox, Mutatiosessor transf |
| 10750 | A | G | RNA | MT-ND4L | NU4LM | P03901 | 94 | N | S | PROVEAN, PhDSNP, MutationTaster, Condel |
| 10775 | G | A | RNA | MT-ND4 | NU4M | P03905 | 6 | V | I | Condel, PolyPhen2 transf, SIFT transf |
| 10808 | C | T | RNA | MT-ND4 | NU4M | P03905 | 17 | L | F | CADD, Condel, MtoolBox |
| 10866 | T | C | RNA | MT-ND4 | NU4M | P03905 | 36 | I | T | Condel |
| 10972 | A | C | RNA | MT-ND4 | NU4M | P03905 | 71 | W | C | PolyPhen2, SIFT, FatHmm, PROVEAN, Mutatiosessor, EFIN_SP, EFIN_HD, CADD, PhDSNP, SP, SNPDryad, COVEC, WMV, MtoolBox, APOGEE, Mutatiosessor transf |
| 11129 | A | G | RNA | MT-ND4 | NU4M | P03905 | 124 | T | A | Condel, SIFT transf |
| 11963 | G | A | RNA | MT-ND4 | NU4M | P03905 | 402 | V | I | Condel, PolyPhen2 transf, SIFT transf |
| 12341 | C | A | RNA | MT-ND5 | NU5M | P03915 | 2 | T | N | Condel, PolyPhen2 transf |
| 12346 | C | T | RNA | MT-ND5 | NU5M | P03915 | 4 | H | Y | Condel, PolyPhen2 transf, SIFT transf |
| 12368 | C | T | RNA | MT-ND5 | NU5M | P03915 | 11 | T | I | FatHmm, PROVEAN, Condel, PolyPhen2 transf |
| 12403 | C | T | RNA | MT-ND5 | NU5M | P03915 | 23 | L | F | Condel, PolyPhen2 transf |
| 12802 | A | G | RNA | MT-ND5 | NU5M | P03915 | 156 | S | G | Condel |
| 13034 | T | C | RNA | MT-ND5 | NU5M | P03915 | 233 | L | P | PolyPhen2, FatHmm, PROVEAN, Mutatiosessor, EFIN_SP, EFIN_HD, CADD, PANTHER, PhDSNP, SP, MutationTaster, SNPDryad, COVEC, WMV, MtoolBox, APOGEE, Mutatiosessor transf |
| 13118 | T | A | RNA | MT-ND5 | NU5M | P03915 | 261 | I | N | PolyPhen2, FatHmm, PROVEAN, Mutatiosessor, EFIN_SP, CADD, PANTHER, PhDSNP, SP, SNPDryad, COVEC, WMV, MtoolBox, APOGEE, Mutatiosessor transf |
| 13145 | G | A | RNA | MT-ND5 | NU5M | P03915 | 270 | S | N | Condel |
| 13226 | A | C | RNA | MT-ND5 | NU5M | P03915 | 297 | D | A | PolyPhen2, FatHmm, PROVEAN, Mutatiosessor, EFIN_SP, CADD, PANTHER, PhDSNP, SP, MutationTaster, SNPDryad, COVEC, WMV, MtoolBox, APOGEE, Mutatiosessor transf |
| 13303 | C | T | RNA | MT-ND5 | NU5M | P03915 | 323 | H | Y | PANTHER, PhDSNP, Condel, SIFT transf |
| 13466 | G | A | RNA | MT-ND5 | NU5M | P03915 | 377 | S | N | FatHmm, PANTHER, Condel |
| 13476 | A | T | RNA | MT-ND5 | NU5M | P03915 | 380 | L | F | PolyPhen2, FatHmm, PROVEAN, Mutatiosessor, EFIN_SP, EFIN_HD, CADD, PANTHER, PhDSNP, SP, SNPDryad, COVEC, WMV, MtoolBox, APOGEE, Mutatiosessor transf |
| 13517 | A | T | RNA | MT-ND5 | NU5M | P03915 | 394 | H | L | PANTHER, Condel, PolyPhen2 transf |
| 13600 | T | A | RNA | MT-ND5 | NU5M | P03915 | 422 | Y | N | PolyPhen2, FatHmmW, FatHmm, PROVEAN, Mutatiosessor, EFIN_SP, CADD, PANTHER, PhDSNP, SP, SNPDryad, COVEC, WMV, MtoolBox, APOGEE, DEOGEN2, Mutatiosessor transf |
| 13614 | A | T | RNA | MT-ND5 | NU5M | P03915 | 426 | M | I | Condel |
| 13618 | C | T | RNA | MT-ND5 | NU5M | P03915 | 428 | L | F | Condel, MtoolBox |
| 13651 | A | G | RNA | MT-ND5 | NU5M | P03915 | 439 | T | A | CADD, Condel, MtoolBox |
| 13763 | C | A | RNA | MT-ND5 | NU5M | P03915 | 476 | S | Y | PolyPhen2, PROVEAN, CADD, PANTHER, SP, Condel, MtoolBox, SIFT transf |
| 13804 | G | A | RNA | MT-ND5 | NU5M | P03915 | 490 | A | T | PolyPhen2, PROVEAN, CADD, PhDSNP, SP, SNPDryad, COVEC, WMV, MtoolBox |
| 13816 | A | C | RNA | MT-ND5 | NU5M | P03915 | 494 | T | P | PolyPhen2, FatHmm, PROVEAN, Mutatiosessor, CADD, PANTHER, PhDSNP, SP, SNPDryad, COVEC, WMV, MtoolBox, Mutatiosessor transf |
| 13886 | T | C | RNA | MT-ND5 | NU5M | P03915 | 517 | L | P | PhDSNP, Condel |
| 14170 | A | T | RNA | MT-ND6 | NU6M | P03923 | 168 | I | M | PolyPhen2, MtoolBox |
| 14172 | T | G | RNA | MT-ND6 | NU6M | P03923 | 168 | I | L | PolyPhen2, CADD, MtoolBox |
| 14280 | A | G | RNA | MT-ND6 | NU6M | P03923 | 132 | S | P | FatHmm, PhDSNP, Condel |

| | | | | | | | | | | |
|-------|---|---|-----|--------|------|--------|-----|---|---|---|
| 14562 | C | T | RNA | MT-ND6 | NU6M | P03923 | 38 | V | I | Condel |
| 14757 | T | C | RNA | MT-CYB | CYB | P00156 | 4 | M | T | Condel, PolyPhen2 transf |
| 14793 | A | G | RNA | MT-CYB | CYB | P00156 | 16 | H | R | PhDSNP, SP, Condel, PolyPhen2 transf |
| 14798 | T | C | RNA | MT-CYB | CYB | P00156 | 18 | F | L | Condel, PolyPhen2 transf |
| 14871 | T | C | RNA | MT-CYB | CYB | P00156 | 42 | I | T | PROVEAN, PhDSNP, Condel, PolyPhen2 transf |
| 14873 | C | A | RNA | MT-CYB | CYB | P00156 | 43 | L | I | Condel |
| 14881 | C | G | RNA | MT-CYB | CYB | P00156 | 45 | I | M | PolyPhen2, FatHmm, PANTHER, PhDSNP, SP, MtoolBox |
| 15077 | G | A | RNA | MT-CYB | CYB | P00156 | 111 | E | K | PolyPhen2, PROVEAN, CADD, PhDSNP, SNPDryad, MtoolBox |
| 15596 | G | A | RNA | MT-CYB | CYB | P00156 | 284 | V | I | Condel |
| 15617 | G | A | RNA | MT-CYB | CYB | P00156 | 291 | V | I | PolyPhen2, CADD, PhDSNP, SP, COVEC, WMV, MtoolBox |
| 15664 | C | A | RNA | MT-CYB | CYB | P00156 | 306 | I | M | CADD, Condel |
| 15699 | G | A | RNA | MT-CYB | CYB | P00156 | 318 | R | H | PolyPhen2, FatHmm, PROVEAN, Mutatiossessor, EFIN_HD, CADD, PANTHER, PhDSNP, SNPDryad, COVEC, WMV, MtoolBox, Mutatiossessor transf |
| 15725 | C | T | RNA | MT-CYB | CYB | P00156 | 327 | L | F | PANTHER, Condel |
| 15734 | G | A | RNA | MT-CYB | CYB | P00156 | 330 | A | T | PhDSNP, Condel |
| 15777 | G | A | RNA | MT-CYB | CYB | P00156 | 344 | S | N | CADD, PhDSNP, Condel |
| 15812 | G | A | RNA | MT-CYB | CYB | P00156 | 356 | V | M | EFIN_SP, MutationTaster, Condel |

Among these 712 variants, 70 of them were associated to diseases (Supplementary Table S9). The 70 variants include 40 protein modifier variants. Among them, 2 were associated to Left Ventricular Non-Compaction disease and 2 to hypertrophic cardiomyopathy.

4. Discussion

We have found that mtDNA copy number in heart tissue was significantly lower in CCC than in healthy individuals, while blood mtDNA content was similar among asymptomatic Chagas disease, moderate and severe CCC patients. While mitochondrial haplogroup lineages of severe CCC patients and healthy heart donors were predominantly African, the percentage of African lineage was much higher among Chagas disease patients from the peripheral blood cohort than healthy individuals from the census-based São Paulo reference population. Conversely, the percentage of the European lineage was under-represented among the Chagas disease cohort than the reference population. The African lineage was over represented in the asymptomatic subgroup when compared with the CCC patient subgroup. Regarding European major haplogroups, H haplogroup was more prevalent among Chagas patients than in the reference population. Whole mtDNA sequencing disclosed that 84 CCC-specific variants were non-synonymous pathogenic changes, some of them previously associated with heart disease.

Here, we have shown that heart from end stage CCC patients display lower mitochondrial DNA copy number than heart tissues from organ donors having no suitable recipient. This result was also consistent with our previous findings that CCC showed even lower mtDNA copy number than other non-inflammatory cardiomyopathies (33) and heart proteomic analysis done on myocardial tissue (30-32, 45). The finding that mtDNA copy number in blood is comparable among Chagas disease clinical groups indicates that the mtDNA copy number variation is tissue specific, and therefore mitochondrial dysfunction detected in CCC hearts is tissue-specific. Indeed, mtDNA copy number variation is depending on cell-type (46). mtDNA copy number variations have been associated to multiple diseases including cardiovascular diseases such as Myocardial Infarction (47), cardiomyopathy, heart failure, and arrhythmias (47). However, our data indicate that blood mtDNA copy number is not a useful clinical biomarker for CCC.

Cardiac mitochondria provide the most ATP in the heart. Therefore, mitochondrial dysfunction is pivotal in heart diseases because of energy supply shortage and increased production of reactive oxidative species (ROS). So, maintaining optimal mitochondrial homeostasis is essential for cardiomyocytes. The total mitochondrial content is mainly controlled by biogenesis and mitochondria-specific autophagy for degradation (48). Damaged mitochondria are normally removed by autophagy/lysosome machinery (49). Several studies described the importance of autophagy in parasite invasion (50). The most characterized mechanism of autophagy is the ATG5-, ATG7-, and LC3-mediated conventional autophagy pathway (51), recent studies have identified an Atg5/Atg7-independent pathway called alternative autophagy (52). In chronic Chagas patients, autophagy activity was not completely investigated. However, based on bulk RNA-seq we investigated the genetic deregulations present in the moderate and severe stages of CCC. Several genes, associated to mitophagy, are differentially expressed in CCC bulk tissues (such as ATG7, LC3, FIP200 and BECN1) (53). Severe CCC is associated to a myocardial inflammatory infiltrate and to an IFN γ /TNF α -rich Th1 type response (22-28). This exacerbated response will lead to ROS/NOS overproduction triggering mitochondrial damage and cardiomyocyte apoptosis. Defective mitophagy or autophagy may be associated to mtDNA release into the cytoplasm, outside

of the cell, or into circulation. Circulating mtDNA enhances expression of type I interferons, and by NLRP3 inflammasome, which promotes the activation of pro-inflammatory cytokines Interleukin-1beta and Interleukin-18 (54). These trigger inflammatory responses in cardiomyocytes and may induce myocarditis, and dilated cardiomyopathy (55).

MtDNA haplogrouping study indicates that African haplogroups were over represented in the Chagas subject groups in compared to healthy individuals. Even if it is not statistically significant, the L0 African sub-haplogroup is more frequent in ASY individuals than in CCC patient groups. The L3 African sub-haplogroup is less frequent in ASY individuals than in the CCC patients. This sub-haplogroup distribution among the groups is consistent with the association of RNR1-921T/C which is a key marker of the L3 African sub-haplogroup. We cannot conclude that the African haplogroup is directly associated with an increased risk of developing severe chronic chagasic cardiomyopathy. It can be simply stated that individuals with African ancestry have a greater chance of lower socioeconomic status that exposes them to parasitic diseases such as Chagas disease. Therefore, the percentage of patients with CCC in the African ethnic group increases. Conversely, the reverse is observed for the European haplogroup. Our results suggest that European macro haplogroup H is associated with increased risk towards CCC. This result agrees with previous work on dilated and hypertrophic cardiomyopathy (56-58). Significantly, intrinsic mitochondrial function has been reported in skeletal muscle myocytes from individuals carrying haplogroup H as compared to haplogroup U (18). This may imply that different haplogroups display distinct mitochondrial function profiles. In the case of Haplogroup H, it is possible to hypothesize that increased OXPHOS is associated with increased ROS production, which may lead to mitochondrial dysfunction and heart damage in CCC patients.

The two main types of mitochondrial DNA mutation are large deletion and point mutations, mostly maternally transmitted, with some de novo mutations have also been reported. Here, no deletion was associated to Chagas disease. The finding of 84 mtDNA-encoded protein sequence pathogenic variants associated with CCC may indicate a role of mtDNA variants in CCC pathogenesis. At this level, the genetic contribution of each of them to the phenotype is undefined. It has been reported that mtDNA variants affect the expressivity of nDNA mutations, leading to experimental cardiomyopathy (59). The functional characterization is still an important bottleneck in mitochondrial diseases. Indeed, in the databases on mitochondrial variants, a very small proportion of these variants have been functionally characterized. Thus, for the vast majority of variants, indications of their functional implications remain putative. Moreover, as no clear variant is associated to disease here, may suggested an additive effect of the mt DNA variants. These studies are complicated by the fact that these variants have been shown to have incomplete penetrance (60). In addition, to complicate matters, mitochondrial disorders are clinically heterogeneous diseases associated with germline mutations in mitochondrial DNA (mtDNA) and/or nuclear DNA-encoded (nDNA) genes, with impaired mitochondrial structure and function. Therefore, mutations at the mtDNA or nDNA levels can have a variety of pathogenesis outcomes. We can thus postulate a synergic effect between these mtDNA variants and nuclear DNA mitochondrial variants previously associated to CCC (36). The description and the characterization of these variants is essential to define a susceptibility polygenic score analysis including all the associated variants.

When the proportion of harmful mtDNA variants reaches a critical level of heteroplasmy, the defects that can give rise to disease. The description of associations of pathogenic variants to multiple diseases, raise the concept of

mitochondria genome editing technology. Gene editing technologies, commonly used for nuclear variants, are not so obvious to apply to mitochondria variants. The first approach leads to decrease the amount of variant mtDNA in mitochondria by specifically targeting and cleaving the mutant mtDNA molecules. Double-strand breaks in mutated mtDNA induce the rapid degradation of the linearized molecule. The residual mtDNA, mostly wild-type, replicates and the heteroplasmy level decreases. The critical point in this approach is the efficiency of the nucleases (mitoREs, mtZFNs, mitoTALENs) (61-63). The development of mitochondria targeted CRISPR/Cas9 systems has been hampered by the lack of efficiency to import guide RNA (gRNA) into the mitochondrial matrix (64, 65). The second approach does not try to decrease the heteroplasmy level but simply to correct the mutated base (66). This concept has been recently described by Kar et al. (67).

There is no cure available for mitochondrial diseases owing to the different genes and phenotypes associated with the cause of such disorders. Nevertheless, few symptomatic treatments have been proven by clinical trials as palliative therapies in the last decade. A mitochondrial cocktail, i.e. a combination of vitamins, cofactors, nutrients and antioxidants, may alleviate symptoms, limit disease progression, and overcome mitochondrial toxins. In the short term, specific dietary restrictions have been shown to ameliorate mitochondrial health in patients with mitochondrial disorders. Thus, high-carbohydrate diet increases oxidative stress, perhaps for individuals with impaired oxidative phosphorylation (68). The ketogenic diet (high-fat diet) has been shown to be beneficial for patients with pyruvate dehydrogenase deficiency, but not for pyruvate carboxylase deficiency and in the treatment of fatty acid oxidation disorders (69). Individuals with disorders of the respiratory chain could be treated with agents that enhance electron transport and substrate delivery and bypass of its components. For example, CoQ10 supplementation and have been shown to reduce high lactate levels after exercise and increase oxygen consumption (70, 71). A diet rich in vitamins and amino acids can be used as a source of redox agents and intracellular buffers for ATP (72). For stroke-like episodes, myopathy, diabetes and lactic acidosis a treatment using natural NO precursors (e.g. arginine and citrulline) was found to restore NO production (68, 73, 74).

Our results are consistent with our previous data based on proteomic, transcriptomic and genetic analyses (26, 31-33, 36, 53, 75, 76). Mitochondrial dysfunction is clearly related to the compromise of mitochondria's ability to make appropriate levels of ATP and to an enhanced formation of reactive oxygen species (ROS). This dysfunction may be the result of mtDNA or nDNA variants, but may also occur as a response to aging and various disease and environmental stresses, leading to the development of cardiomyopathies (CMs), ROS accumulation and cell death. Mitochondrial dysfunction can have a direct deleterious effect on the contractile capacity of myocardial cells. Defects in mitochondrial regulation of calcium homeostasis can alter mechanical function and electrical conduction (77). Moreover, pathogenic inflammatory responses produced by damaged mitochondria can particularly trigger the atherosclerotic process. ROS generated in the mitochondria may lead to target-organ damage, dysfunction, hypertrophy, and inflammation. Hypertension and cardiac fibrosis are associated with depletion and inactivation of the key mitochondrial deacetylase, sirtuin 3, which is involved in the control of key metabolic steps (78-80).

The finding that pathogenic mitochondrial protein-coding SNPs and mitochondrial haplogroups associate with risk of evolving to CCC is consistent with a key role of mitochondrial DNA in the development of Chronic Chagas disease Cardiomyopathy and are in line with previous reports of nuclear

DNA-encoded mitochondrial gene variants. These may interact and increase mitochondrial susceptibility to IFN- γ -induced damage in the CCC myocardium, leading to ventricular dysfunction.

5. Conclusions

Cardiomyopathies are major causes of heart failure. Chagas disease (CD) is caused by the parasite *Trypanosoma cruzi*, and it is endemic in Central, South America. Here, we described that mtDNA copy number was significantly lower in CCC than in heart tissue from healthy individuals, while blood mtDNA content was similar among asymptomatic Chagas disease, moderate and severe CCC patients. The mitochondria whereas the European lineage is associated to protection against cardiomyopathy. Finally, 84 mtDNA-encoded protein sequence pathogenic variants were associated with CCC.

Supplementary Materials: The following supporting information can be downloaded at: www.mdpi.com/xxx/s1, Table S1: Heart tissue samples used for mitochondria quantification. Table S2: Blood DNAs samples used for mitochondria quantification. Table S3: Blood samples used for haplogroup analysis. Table S4: Statistics on African sub-haplogroup distribution. Table S5: Clinical information for patients whom the whole mitochondria genome has been sequenced (n=112). Table S6: Carrier analysis performed on 12429 Variants. Table S7: Heteroplasmy level test between the three groups (ASY, moderate CCC and severe CCC) and heteroplasmy level analyses between the groups (two by two: ASY vs CCC; ASY vs Moderate CCC; ASY vs Severe CCC and Moderate CCC vs Severe CCC). Table S8: Variant annotation. Table S9: List of the variants associated to diseases.

Author Contributions: Experimental work: FG, JPSN, PA; Bioinformatic analysis: Pauline Brochet, DG; Biological sample collect: BMI, CM, AFF, RHBS, AK, Samuel Steffen, ANS, PP, AIF, EAB, CWP, BS, FCD, MFS, FAG, JAMN, AF, RCFZ, Sergio Siqueira, GDLP, FB, Paula Buck, RRA, HTLW, AS, MHH, EAD, ACP, VRJ, MM; Writing—original draft preparation: Pauline Brochet, PA, JPSN, EC-N, CC; Writing—review and editing: MN, JK, VP, EC-N, CC.

Funding: This work was supported by the Institut National de la Santé et de la Recherche Médicale (INSERM); the Aix-Marseille University (grant number: AMIDEX “International_2018” MITOMUTCHAGAS); the French Agency for Research (Agence Nationale de la Recherche-ANR (grant numbers: “Br-Fr-Chagas”, “landscardio”). This work was founded by the Inserm Cross-Cutting Project GOLD. This project has received funding from the Ex-cellecne Initiative of Aix-Marseille University - A*Midex a French “Investissements d’Avenir programme”- Institute MarMaRa AMX-19-IET-007. This work was supported by FAPESP (São Paulo State Research Funding Agency Brazil) (grant 2022/00758-0). Brazilian National Research Council (CNPq) provided a Productivity Award for ECN and JK. This work is supported by the CAPES-COFECUB program (Me987/23). CC is supported by the Inserm’s PRI/IRP 2022 program (IFNG-Mito). The funders did not play any role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Institutional Review Board Statement: The protocol was approved by the institutional review boards of the University of São Paulo School of Medicine (n°009/2011, January 21th 2011) and INSERM (French National Institute of Health and Medical Research) (FWA00005831, March 16th 2011). All experimental methods comply with the Helsinki Declaration.

Informed Consent Statement: Written informed consent was obtained from all patients.

Data Availability Statement: The data analyzed during the current study is available upon reasonable requests to the corresponding authors.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Richardson P, McKenna W, Bristow M, Maisch B, Mautner B, O'Connell J, Olsen E, Thiene G, Goodwin J, Gyarfás I, et al. Report of the 1995 World Health Organization/International Society and Federation of Cardiology Task Force on the Definition and Classification of cardiomyopathies. *Circulation*. 1996;93(5):841-2.
2. Bozkurt B, Colvin M, Cook J, Cooper LT, Deswal A, Fonarow GC, Francis GS, Lenihan D, Lewis EF, McNamara DM, et al. Current Diagnostic and Treatment Strategies for Specific Dilated Cardiomyopathies: A Scientific Statement From the American Heart Association. *Circulation*. 2016;134(23):e579-e646.
3. Mestroni L, Brun F, Spezzacatene A, Sinagra G, Taylor MR. Genetic Causes of Dilated Cardiomyopathy. *Prog Pediatr Cardiol*. 2014;37(1-2):13-8.
4. Bakalakos A, Ritsatos K, Anastasakis A. Current perspectives on the diagnosis and management of dilated cardiomyopathy Beyond heart failure: a Cardiomyopathy Clinic Doctor's point of view. *Hellenic J Cardiol*. 2018;59(5):254-61.
5. Akhtar MM, Lorenzini M, Cicerchia M, Ochoa JP, Hey TM, Sabater Molina M, Restrepo-Cordoba MA, Dal Ferro M, Stolfo D, Johnson R, et al. Clinical Phenotypes and Prognosis of Dilated Cardiomyopathy Caused by Truncating Variants in the TTN Gene. *Circ Heart Fail*. 2020;13(10):e006832.
6. Lee JH, Lee SE, Cho MC. Clinical Implication of Genetic Testing in Dilated Cardiomyopathy. *Int J Heart Fail*. 2022;4(1):1-11.
7. Lu JT, Muchir A, Nagy PL, Worman HJ. LMNA cardiomyopathy: cell biology and genetics meet clinical medicine. *Dis Model Mech*. 2011;4(5):562-8.
8. Tabish AM, Azzimato V, Alexiadis A, Buyandelger B, Knoll R. Genetic epidemiology of titin-truncating variants in the etiology of dilated cardiomyopathy. *Biophys Rev*. 2017;9(3):207-23.
9. Zak R, Rabinowitz M, Rajamanickam C, Merten S, Kwiatkowska-Patzer B. Mitochondrial proliferation in cardiac hypertrophy. *Basic Res Cardiol*. 1980;75(1):171-8.
10. Goffart S, von Kleist-Retzow JC, Wiesner RJ. Regulation of mitochondrial proliferation in the heart: power-plant failure contributes to cardiac failure in hypertrophy. *Cardiovasc Res*. 2004;64(2):198-207.
11. El-Hattab AW, Scaglia F. Mitochondrial Cardiomyopathies. *Front Cardiovasc Med*. 2016;3:25.
12. Potgieter M, Pretorius E, Pepper MS. Primary and secondary coenzyme Q10 deficiency: the role of therapeutic supplementation. *Nutr Rev*. 2013;71(3):180-8.
13. Thorburn DR, Sugiana C, Salemi R, Kirby DM, Worgan L, Ohtake A, Ryan MT. Biochemical and molecular diagnosis of mitochondrial respiratory chain disorders. *Biochim Biophys Acta*. 2004;1659(2-3):121-8.
14. Fillmore N, Mori J, Lopaschuk GD. Mitochondrial fatty acid oxidation alterations in heart failure, ischaemic heart disease and diabetic cardiomyopathy. *Br J Pharmacol*. 2014;171(8):2080-90.
15. Nickel A, Kohlhaas M, Maack C. Mitochondrial reactive oxygen species production and elimination. *J Mol Cell Cardiol*. 2014;73:26-33.
16. Richter-Dennerlein R, Korwitz A, Haag M, Tatsuta T, Dargazanli S, Baker M, Decker T, Lamkemeyer T, Rugarli EL, Langer T. DNAJC19, a mitochondrial cochaperone associated with cardiomyopathy, forms a complex with prohibitins to regulate cardiolipin remodeling. *Cell Metab*. 2014;20(1):158-71.
17. Kohda M, Tokuzawa Y, Kishita Y, Nyuzuki H, Moriyama Y, Mizuno Y, Hirata T, Yatsuka Y, Yamashita-Sugahara Y, Nakachi Y, et al. A Comprehensive Genomic Analysis Reveals the Genetic Landscape of Mitochondrial Respiratory Chain Complex Deficiencies. *PLoS Genet*. 2016;12(1):e1005679.
18. Larsen S, Diez-Sanchez C, Rabol R, Ara I, Dela F, Helge JW. Increased intrinsic mitochondrial function in humans with mitochondrial haplogroup H. *Biochim Biophys Acta*. 2014;1837(2):226-31.
19. Perez-Molina JA, Molina I. Chagas disease. *Lancet*. 2018;391(10115):82-94.
20. Morillo CA, Marin-Neto JA, Avezum A, Sosa-Estani S, Rassi A, Jr., Rosas F, Villena E, Quiroz R, Bonilla R, Britto C, et al. Randomized Trial of Benznidazole for Chronic Chagas' Cardiomyopathy. *N Engl J Med*. 2015;373(14):1295-306.

21. Bestetti RB, Muccillo G. Clinical course of Chagas' heart disease: a comparison with dilated cardiomyopathy. *Int J Cardiol.* 1997;60(2):187-93. 681 682
22. Mady C, Cardoso RH, Barretto AC, da Luz PL, Bellotti G, Pileggi F. Survival and predictors of survival in patients with congestive heart failure due to Chagas' cardiomyopathy. *Circulation.* 1994;90(6):3098-102. 683 684 685
23. Higuchi ML, De Moraes CF, Pereira Barreto AC, Lopes EA, Stolf N, Bellotti G, Pileggi F. The role of active myocarditis in the development of heart failure in chronic Chagas' disease: a study based on endomyocardial biopsies. *Clin Cardiol.* 1987;10(11):665-70. 686 687 688
24. Pereira Barreto AC, Mady C, Arteaga-Fernandez E, Stolf N, Lopes EA, Higuchi ML, Bellotti G, Pileggi F. Right ventricular endomyocardial biopsy in chronic Chagas' disease. *Am Heart J.* 1986;111(2):307-12. 689 690 691
25. Abel LC, Rizzo LV, Ianni B, Albuquerque F, Bacal F, Carrara D, Bocchi EA, Teixeira HC, Mady C, Kalil J, et al. Chronic Chagas' disease cardiomyopathy patients display an increased IFN-gamma response to *Trypanosoma cruzi* infection. *J Autoimmun.* 2001;17(1):99-107. 692 693 694
26. Nogueira LG, Santos RH, Ianni BM, Fiorelli AI, Mairena EC, Benvenuti LA, Frade A, Donadi E, Dias F, Saba B, et al. Myocardial chemokine expression and intensity of myocarditis in Chagas cardiomyopathy are controlled by polymorphisms in CXCL9 and CXCL10. *PLoS Negl Trop Dis.* 2012;6(10):e1867. 695 696 697
27. Talvani A, Rocha MO, Barcelos LS, Gomes YM, Ribeiro AL, Teixeira MM. Elevated concentrations of CCL2 and tumor necrosis factor-alpha in chagasic cardiomyopathy. *Clin Infect Dis.* 2004;38(7):943-50. 698 699
28. Teixeira MM, Gazzinelli RT, Silva JS. Chemokines, inflammation and *Trypanosoma cruzi* infection. *Trends Parasitol.* 2002;18(6):262-5. 700 701
29. Laugier L, Ferreira LRP, Ferreira FM, Cabantous S, Frade AF, Nunes JP, Ribeiro RA, Brochet P, Teixeira PC, Santos RHB, et al. miRNAs may play a major role in the control of gene expression in key pathobiological processes in Chagas disease cardiomyopathy. *PLoS Negl Trop Dis.* 2020;14(12):e0008889. 702 703 704
30. Cunha-Neto E, Teixeira PC, Fonseca SG, Bilate AM, Kalil J. Myocardial gene and protein expression profiles after autoimmune injury in Chagas' disease cardiomyopathy. *Autoimmun Rev.* 2011;10(3):163-5. 705 706
31. Teixeira PC, Ducret A, Langen H, Nogoceke E, Santos RHB, Silva Nunes JP, Benvenuti L, Levy D, Bydlowski SP, Bocchi EA, et al. Impairment of Multiple Mitochondrial Energy Metabolism Pathways in the Heart of Chagas Disease Cardiomyopathy Patients. *Front Immunol.* 2021;12:755782. 707 708 709
32. Teixeira PC, Santos RH, Fiorelli AI, Bilate AM, Benvenuti LA, Stolf NA, Kalil J, Cunha-Neto E. Selective decrease of components of the creatine kinase system and ATP synthase complex in chronic Chagas disease cardiomyopathy. *PLoS Negl Trop Dis.* 2011;5(6):e1205. 710 711 712
33. Nunes JPS, Andrieux P, Brochet P, Almeida RR, Kitano E, Honda AK, Iwai LK, Andrade-Silva D, Goudenege D, Alcantara Silva KD, et al. Co-Exposure of Cardiomyocytes to IFN-gamma and TNF-alpha Induces Mitochondrial Dysfunction and Nitro-Oxidative Stress: Implications for the Pathogenesis of Chronic Chagas Disease Cardiomyopathy. *Front Immunol.* 2021;12:755862. 713 714 715 716
34. Kolwicz SC, Jr., Purohit S, Tian R. Cardiac metabolism and its interactions with contraction, growth, and survival of cardiomyocytes. *Circ Res.* 2013;113(5):603-16. 717 718
35. Zicker F, Smith PG, Netto JC, Oliveira RM, Zicker EM. Physical activity, opportunity for reinfection, and sibling history of heart disease as risk factors for Chagas' cardiopathy. *Am J Trop Med Hyg.* 1990;43(5):498-505. 719 720 721
36. Ouarhache M, Marquet S, Frade AF, Ferreira AM, Ianni B, Almeida RR, Nunes JPS, Ferreira LRP, Rigaud VO, Candido D, et al. Rare Pathogenic Variants in Mitochondrial and Inflammation-Associated Genes May Lead to Inflammatory Cardiomyopathy in Chagas Disease. *J Clin Immunol.* 2021;41(5):1048-63. 722 723 724
37. Quiros PM, Goyal A, Jha P, Auwerx J. Analysis of mtDNA/nDNA Ratio in Mice. *Curr Protoc Mouse Biol.* 2017;7(1):47-54. 725 726
38. Naslavsky MS, Scliar MO, Yamamoto GL, Wang JYT, Zverinova S, Karp T, Nunes K, Ceroni JRM, de Carvalho DL, da Silva Simoes CE, et al. Whole-genome sequencing of 1,171 elderly admixed individuals from Sao Paulo, Brazil. *Nat Commun.* 2022;13(1):1004. 727 728 729
39. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2010;26(5):589-95. 730 731

40. Herzeel C, Costanza P, Decap D, Fostier J, Verachtert W. elPrep 4: A multithreaded framework for sequence analysis. *PLoS One*. 2019;14(2):e0209523. 732 733
41. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet*. 2011;43(5):491-8. 734 735 736
42. Kim S, Scheffler K, Halpern AL, Bekritsky MA, Noh E, Kallberg M, Chen X, Kim Y, Beyter D, Krusche P, et al. Strelka2: fast and accurate calling of germline and somatic variants. *Nat Methods*. 2018;15(8):591-4. 737 738 739
43. Goudenege D, Bris C, Hoffmann V, Desquirit-Dumas V, Jardel C, Rucheton B, Bannwarth S, Paquis-Flucklinger V, Lebre AS, Colin E, et al. eKLIPse: a sensitive tool for the detection and quantification of mitochondrial DNA deletions from next-generation sequencing data. *Genet Med*. 2019;21(6):1407-16. 740 741 742
44. Castellana S, Ronai J, Mazza T. MitImpact: an exhaustive collection of pre-computed pathogenicity predictions of human mitochondrial non-synonymous variants. *Hum Mutat*. 2015;36(2):E2413-22. 743 744
45. Teixeira PC, Iwai LK, Kuramoto AC, Honorato R, Fiorelli A, Stolf N, Kalil J, Cunha-Neto E. Proteomic inventory of myocardial proteins from patients with chronic Chagas' cardiomyopathy. *Braz J Med Biol Res*. 2006;39(12):1549-62. 745 746 747
46. Picard M. Blood mitochondrial DNA copy number: What are we counting? *Mitochondrion*. 2021;60:1-11. 748 749
47. Sundquist K, Sundquist J, Palmer K, Memon AA. Role of mitochondrial DNA copy number in incident cardiovascular diseases and the association between cardiovascular disease and type 2 diabetes: A follow-up study on middle-aged women. *Atherosclerosis*. 2022;341:58-62. 750 751 752
48. Nah J. The Role of Alternative Mitophagy in Heart Disease. *Int J Mol Sci*. 2023;24(7). 753
49. Dutta D, Calvani R, Bernabei R, Leeuwenburgh C, Marzetti E. Contribution of impaired mitochondrial autophagy to cardiac aging: mechanisms and therapeutic opportunities. *Circ Res*. 2012;110(8):1125-38. 754 755 756
50. Romano PS, Cueto JA, Casassa AF, Vanrell MC, Gottlieb RA, Colombo MI. Molecular and cellular mechanisms involved in the *Trypanosoma cruzi*/host cell interplay. *IUBMB Life*. 2012;64(5):387-96. 757 758
51. Yin Z, Pascual C, Klionsky DJ. Autophagy: machinery and regulation. *Microb Cell*. 2016;3(12):588-96. 759 760
52. Nishida Y, Arakawa S, Fujitani K, Yamaguchi H, Mizuta T, Kanaseki T, Komatsu M, Otsu K, Tsujimoto Y, Shimizu S. Discovery of Atg5/Atg7-independent alternative macroautophagy. *Nature*. 2009;461(7264):654-8. 761 762 763
53. Brochet P, Ianni BM, Laugier L, Frade AF, Silva Nunes JP, Teixeira PC, Mady C, Ferreira LRP, Ferre Q, Santos RHB, et al. Epigenetic regulation of transcription factor binding motifs promotes Th1 response in Chagas disease cardiomyopathy. *Front Immunol*. 2022;13:958200. 764 765 766
54. De Gaetano A, Solodka K, Zanini G, Selleri V, Mattioli AV, Nasi M, Pinti M. Molecular Mechanisms of mtDNA-Mediated Inflammation. *Cells*. 2021;10(11). 767 768
55. Oka T, Hikoso S, Yamaguchi O, Taneike M, Takeda T, Tamai T, Oyabu J, Murakawa T, Nakayama H, Nishida K, et al. Mitochondrial DNA that escapes from autophagy causes inflammation and heart failure. *Nature*. 2012;485(7397):251-5. 769 770 771
56. Hagen CM, Aidt FH, Hedley PL, Jensen MK, Havndrup O, Kanters JK, Moolman-Smook JC, Larsen SO, Bundgaard H, Christiansen M. Mitochondrial haplogroups modify the risk of developing hypertrophic cardiomyopathy in a Danish population. *PLoS One*. 2013;8(8):e71904. 772 773 774
57. Govindaraj P, Rani B, Sundaravadeivel P, Vanniarajan A, Indumathi KP, Khan NA, Dhandapany PS, Rani DS, Tamang R, Bahl A, et al. Mitochondrial genome variations in idiopathic dilated cardiomyopathy. *Mitochondrion*. 2019;48:51-9. 775 776 777
58. Fernandez-Caggiano M, Barallobre-Barreiro J, Rego-Perez I, Crespo-Leiro MG, Paniagua MJ, Grille Z, Blanco FJ, Domenech N. Mitochondrial DNA haplogroup H as a risk factor for idiopathic dilated cardiomyopathy in Spanish population. *Mitochondrion*. 2013;13(4):263-8. 778 779 780
59. McManus MJ, Picard M, Chen HW, De Haas HJ, Potluri P, Leipzig J, Towheed A, Angelin A, Sengupta P, Morrow RM, et al. Mitochondrial DNA Variation Dictates Expressivity and Progression of Nuclear DNA Mutations Causing Cardiomyopathy. *Cell Metab*. 2019;29(1):78-90 e5. 781 782 783

60. Garret P, Bris C, Procaccio V, Amati-Bonneau P, Vabres P, Houcinat N, Tisserant E, Feillet F, Bruel AL, Quere V, et al. Deciphering exome sequencing data: Bringing mitochondrial DNA variants to light. *Hum Mutat.* 2019;40(12):2430-43. 784-786
61. Yahata N, Matsumoto Y, Omi M, Yamamoto N, Hata R. TALEN-mediated shift of mitochondrial DNA heteroplasmy in MELAS-iPSCs with m.13513G>A mutation. *Sci Rep.* 2017;7(1):15557. 787-788
62. Gammage PA, Minczuk M. Enhanced Manipulation of Human Mitochondrial DNA Heteroplasmy In Vitro Using Tunable mtZFN Technology. *Methods Mol Biol.* 2018;1867:43-56. 789-790
63. Srivastava S, Moraes CT. Manipulating mitochondrial DNA heteroplasmy by a mitochondrially targeted restriction endonuclease. *Hum Mol Genet.* 2001;10(26):3093-9. 791-792
64. Hussain SA, Yalvac ME, Khoo B, Eckardt S, McLaughlin KJ. Adapting CRISPR/Cas9 System for Targeting Mitochondrial Genome. *Front Genet.* 2021;12:627050. 793-794
65. Bian WP, Chen YL, Luo JJ, Wang C, Xie SL, Pei DS. Knock-In Strategy for Editing Human and Zebrafish Mitochondrial DNA Using Mito-CRISPR/Cas9 System. *ACS Synth Biol.* 2019;8(4):621-32. 795-796
66. Porto EM, Komor AC, Slaymaker IM, Yeo GW. Base editing: advances and therapeutic opportunities. *Nat Rev Drug Discov.* 2020;19(12):839-59. 797-798
67. Kar B, Castillo SR, Sabharwal A, Clark KJ, Ekker SC. Mitochondrial Base Editing: Recent Advances towards Therapeutic Opportunities. *Int J Mol Sci.* 2023;24(6). 799-800
68. El-Hattab AW, Hsu JW, Emrick LT, Wong LJ, Craigen WJ, Jahoor F, Scaglia F. Restoration of impaired nitric oxide production in MELAS syndrome with citrulline and arginine supplementation. *Mol Genet Metab.* 2012;105(4):607-14. 801-803
69. Bough KJ, Wetherington J, Hassel B, Pare JF, Gawryluk JW, Greene JG, Shaw R, Smith Y, Geiger JD, Dingleline RJ. Mitochondrial biogenesis in the anticonvulsant mechanism of the ketogenic diet. *Ann Neurol.* 2006;60(2):223-35. 804-806
70. Rotig A, Appelkvist EL, Geromel V, Chretien D, Kadhom N, Edery P, Lebideau M, Dallner G, Munnich A, Ernster L, et al. Quinone-responsive multiple respiratory-chain dysfunction due to widespread coenzyme Q10 deficiency. *Lancet.* 2000;356(9227):391-5. 807-809
71. Di Giovanni S, Mirabella M, Spinazzola A, Crociani P, Silvestri G, Broccolini A, Tonali P, Di Mauro S, Servidei S. Coenzyme Q10 reverses pathological phenotype and reduces apoptosis in familial CoQ10 deficiency. *Neurology.* 2001;57(3):515-8. 810-812
72. Parikh S, Saneto R, Falk MJ, Anselm I, Cohen BH, Haas R, Medicine Society TM. A modern approach to the treatment of mitochondrial disease. *Curr Treat Options Neurol.* 2009;11(6):414-30. 813-814
73. Pfeffer G, Majamaa K, Turnbull DM, Thorburn D, Chinnery PF. Treatment for mitochondrial disorders. *Cochrane Database Syst Rev.* 2012;2012(4):CD004426. 815-816
74. Almannai M, El-Hattab AW. Nitric Oxide Deficiency in Mitochondrial Disorders: The Utility of Arginine and Citrulline. *Front Mol Neurosci.* 2021;14:682780. 817-818
75. Andrieux P, Chevillard C, Cunha-Neto E, Nunes JPS. Mitochondria as a Cellular Hub in Infection and Inflammation. *Int J Mol Sci.* 2021;22(21). 819-820
76. Laugier L, Frade AF, Ferreira FM, Baron MA, Teixeira PC, Cabantous S, Ferreira LRP, Louis L, Rigaud VOC, Gaiotto FA, et al. Whole-Genome Cardiac DNA Methylation Fingerprint and Gene Expression Analysis Provide New Insights in the Pathogenesis of Chronic Chagas Disease Cardiomyopathy. *Clin Infect Dis.* 2017;65(7):1103-11. 821-824
77. Lopez-Crisosto C, Pennanen C, Vasquez-Trincado C, Morales PE, Bravo-Sagua R, Quest AFG, Chiong M, Lavandero S. Sarcoplasmic reticulum-mitochondria communication in cardiovascular pathophysiology. *Nat Rev Cardiol.* 2017;14(6):342-60. 825-827
78. Zhang Q, Li D, Dong X, Zhang X, Liu J, Peng L, Meng B, Hua Q, Pei X, Zhao L, et al. LncDACH1 promotes mitochondrial oxidative stress of cardiomyocytes by interacting with sirtuin3 and aggravates diabetic cardiomyopathy. *Sci China Life Sci.* 2022;65(6):1198-212. 828-830
79. Chen WJ, Cheng Y, Li W, Dong XK, Wei JL, Yang CH, Jiang YH. Quercetin Attenuates Cardiac Hypertrophy by Inhibiting Mitochondrial Dysfunction Through SIRT3/PARP-1 Pathway. *Front Pharmacol.* 2021;12:739615. 831-833
80. Chen J, Chen S, Zhang B, Liu J. SIRT3 as a potential therapeutic target for heart failure. *Pharmacol Res.* 2021;165:105432. 834-835

2. Différentes variables épigénétiques ont un impact dans le processus pathogénique associé aux cardiomyopathies chagasiennes chroniques

2.1. Des miARN sont associés aux cardiomyopathies chagasiennes chroniques, et semblent réguler des fonctions biologiques essentielles au développement de cette pathologie

Article 2 : miRNAs may play a major role in the control of gene expression in key pathobiological processes in Chagas disease cardiomyopathy.

Laurie Laugier, Ludmila Rodrigues Pinto Ferreira, Frederico Moraes Ferreira, Sandrine Cabantous, Amanda Farage Frade, João Paulo Nunes, Rafael Almeida Ribeiro, **Pauline Brochet**, Priscila Camillo Teixeira, Ronaldo Honorato Barros Santos, Edimar A Bocchi, Fernando Bacal, Darlan da Silva Cândido, Vanessa Escolano Maso, Helder I Nakaya, Jorge Kalil, Edecio Cunha-Neto*, Christophe Chevillard*. PLoS Negl Trop Dis. 2020 Dec 22;14(12):e0008889. doi: 10.1371/journal.pntd.0008889. PMID: 33351798; PMCID: PMC7787679.

Présentation de la publication :

L'implication des miRNA dans les CCC avait déjà été prouvée chez la souris, mais son impact chez l'homme restait à déterminer. À partir du tissu cardiaque de 10 patients atteints de CCC et de 7 individus sains, l'expression de 754 miARN a été quantifiée par RT-PCR, dont 210 exprimés dans tous les échantillons. Parmi eux, 80 ont été associés aux CCC (valeur absolue du Fold Change $\geq 1,5$ et p-valeur $< 0,05$), notamment miR-146a, miR-155, miR-1, miR-133a et miR-133b. Une Analyse en Composantes Principales (ACP) a été menée à partir de l'expression de ces 80 miARN. Les deux premières composantes suffisent à discriminer les cas contrôles des CCC, représentant plus de 84% de la variance totale.

En parallèle, 1.535 gènes, mesurés avec une puce à ADN, ont été identifiés comme différentiellement exprimés (DE) entre les deux phénotypes. La pertinence de ces résultats a été validée par qPCR, où 38 gènes sur 44 gènes présentant une expression différentielle ont été expérimentalement confirmés. Comme précédemment, une ACP réalisée avec les 1.535 gènes DE a permis de séparer les individus en fonction de leur phénotype, les deux premières dimensions présentant 68% de la variation des données.

Les principaux processus biologiques affectés étaient essentiellement liés à la réponse immunitaire (activation des lymphocytes Th1/Th2, présentation d'antigène, développement des lymphocytes B, phagocytose dans les macrophages, cellules NK...), également associés au stress oxydatif, aux processus mitochondriaux ou aux pathologies cardiaques (contractilité du coeur, arythmie). Parmi les gènes DE, d'importants régulateurs ont été identifiés, tels que l'IFN- γ , le TNF- α , ou les facteurs de transcriptions NF-kB, Ap1 et PI3K.

L'association entre les miARN et leurs potentiels gènes cibles a été faite avec le logiciel Ingenuity Pathway Analysis (IPA), en ne conservant que les miARN et gènes DE présentant un profil d'expression opposé. Au total, 571 interactions ont été prédites, impliquant 67 miARN et 396 gènes. En se focalisant sur les gènes DE liés à des processus biologiques d'intérêt (immunitaires, mitochondriaux et cardiaques), 5 miARN (miR-125b, miR-15a, miR-296, miR-29 et miR-103) semblent cibler plus de 20 gènes chacun, impliqués dans la fibrose, l'hypertrophie, l'arythmie, l'inflammation et la réponse Th1, suggérant un grand impact dans les CCC.

Cette étude démontre l'association des miARN dans les CCC chez l'homme, et leur potentiel impact régulateur dans des fonctions biologiques clé. Ainsi, elle apporte des informations complémentaires sur le processus pathogénique associé à cette maladie.

Contribution personnelle :

Ma contribution dans cette analyse a été mineure. Je me suis principalement occupée des analyses en composantes principales, et j'ai participé à la rédaction et aux corrections de l'article.

RESEARCH ARTICLE

miRNAs may play a major role in the control of gene expression in key pathobiological processes in Chagas disease cardiomyopathy

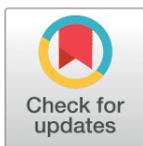
Laurie Laugier¹ , Ludmila Rodrigues Pinto Ferreira^{2,3,4} , Frederico Moraes Ferreira^{2,3,4} , Sandrine Cabantous¹, Amanda Farage Frade^{2,3,4} , Joao Paulo Nunes^{2,3,4} , Rafael Almeida Ribeiro^{2,3,4}, Pauline Brochet⁵ , Priscila Camillo Teixeira^{2,3,4} , Ronaldo Honorato Barros Santos⁶, Edimar A. Bocchi⁶, Fernando Bacal⁶, Darlan da Silva Cândido^{2,3,4}, Vanessa Escolano Maso⁷ , Helder I. Nakaya^{7,8} , Jorge Kalil^{2,3,4} , Edecio Cunha-Neto^{2,3,4} , Christophe Chevillard⁵  ^{5†*}

1 Aix Marseille Université, Génétique et Immunologie des Maladies Parasitaires, Unité Mixte de Recherche S906, Marseille, France; INSERM, U906, Marseille, France, **2** Laboratory of Immunology, Heart Institute (InCor), University of São Paulo, School of Medicine, São Paulo, Brazil, **3** Division of Clinical Immunology and Allergy, University of São Paulo, School of Medicine, São Paulo, Brazil, **4** Institute for Investigation in Immunology (iii), INCT, São Paulo, Brazil, **5** Aix Marseille Université, TAGC Theories and Approaches of Genomic Complexity, Inserm, INSERM, UMR_1090, Marseille, France, **6** Division of Transplantation, Heart Institute (InCor), University of São Paulo, School of Medicine, São Paulo, Brazil, **7** Department of Pathophysiology and Toxicology, School of Pharmaceutical Sciences, University of São Paulo, São Paulo, Brazil, **8** Scientific Platform Pasteur, University of São Paulo, São Paulo, Brazil

 These authors contributed equally to this work.

† ECN and CC also contributed equally to this work.

* edecunha@gmail.com (ECN); christophe.chevillard@univ-amu.fr (CC)



OPEN ACCESS

Citation: Laugier L, Ferreira LRP, Ferreira FM, Cabantous S, Frade AF, Nunes JP, et al. (2020) miRNAs may play a major role in the control of gene expression in key pathobiological processes in Chagas disease cardiomyopathy. PLoS Negl Trop Dis 14(12): e0008889. <https://doi.org/10.1371/journal.pntd.0008889>

Editor: Juan M. Bustamante, University of Georgia, UNITED STATES

Received: March 30, 2020

Accepted: October 14, 2020

Published: December 22, 2020

Copyright: © 2020 Laugier et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Gene expression data were deposited in the GEO database (GSE84796 and GSE111544).

Funding: This work was supported by the Institut National de la Santé et de la Recherche Médicale (INSERM), Aix-Marseille University (Direction des Relations Internationales), the ARCUS II PACA Brésil program, CNPq (the Brazilian National Research Council), and FAPESP (São Paulo State Research Funding Agency-Brazil). ECN and CC

Abstract

Chronic Chagas disease cardiomyopathy (CCC), an especially aggressive inflammatory dilated cardiomyopathy caused by lifelong infection with the protozoan *Trypanosoma cruzi*, is a major cause of cardiomyopathy in Latin America. Although chronic myocarditis may play a major pathogenetic role, little is known about the molecular mechanisms responsible for its severity. The aim of this study is to study the genes and microRNAs expression in tissues and their connections in regards to the pathobiological processes. To do so, we integrated for the first time global microRNA and mRNA expression profiling from myocardial tissue of CCC patients employing pathways and network analyses. We observed an enrichment in biological processes and pathways associated with the immune response and metabolism. IFN γ , TNF and NF κ B were the top upstream regulators. The intersections between differentially expressed microRNAs and differentially expressed target mRNAs showed an enrichment in biological processes such as Inflammation, inflammation, Th1/IFN- γ -inducible genes, fibrosis, hypertrophy, and mitochondrial/oxidative stress/antioxidant response. MicroRNAs also played a role in the regulation of gene expression involved in the key cardiomyopathy-related processes fibrosis, hypertrophy, myocarditis and arrhythmia. Significantly, a discrete number of differentially expressed microRNAs targeted a high number of differentially expressed mRNAs (>20) in multiple processes. Our results suggest that miRNAs orchestrate expression of multiple genes in the major pathophysiological

were recipient for an international program funded either by the French ANR and the Brazilian FAPESP agencies (Br-Fr-chagas). AFF hold fellowships from the São Paulo State Research Funding Agency, FAPESP. ECN and JK have received a Council for Scientific and Technological Development - CNPq productivity award. CC is a recipient of a temporary professorship position supported by the French consulate in Brazil and the University of São Paulo (USP). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

processes in CCC heart tissue. This may have a bearing on pathogenesis, biomarkers and therapy.

Author summary

Chronic Chagas disease cardiomyopathy (CCC), an aggressive dilated cardiomyopathy caused by *Trypanosoma cruzi*, is a major cause of cardiomyopathy in Latin America. Little is known about the molecular mechanisms responsible for its severity. Authors study the possible role of microRNAs in the regulation of gene expression in relevant pathways and pathobiological processes. Differentially expressed genes (DEGs) and differentially expressed miRNAs (DEMs) -small RNAs that can regulate gene expression—associated to severe cardiomyopathy development. The inflammatory mediator Interferon- γ was the most likely inducer of gene expression in CCC, and most genes belonged to the immune response, fibrosis, hypertrophy and mitochondrial metabolism. A discrete number of differentially expressed mRNAs targeted a high number of differentially expressed mRNAs in multiple processes. Moreover, several pathways had multiple targets regulated by microRNAs, suggesting synergic effect. Results suggest that microRNAs orchestrate expression of multiple genes in the major pathophysiological processes in CCC heart tissue.

Introduction

Chagas disease is a major public health problem in Latin America, resulting from lifelong infection with the protozoan parasite *Trypanosoma cruzi*. Up to 30 years after acute infection, approximately 30% of the 6 million infected people eventually develop chronic Chagas cardiomyopathy (CCC), a life-threatening inflammatory dilated cardiomyopathy [1,2]. Most other *T. cruzi*-infected patients will remain asymptomatic for life (60%) or develop digestive disease, which causes less deaths (approx. 10%) [1]. Chagas disease is the most common cause of non-ischemic cardiomyopathy in Latin America, causing approximately 10,000 deaths/year, mainly due to heart failure and severe arrhythmia/sudden death [1]. Migration turned Chagas disease into a global health problem, with an estimated 400,000 infected persons living in nonendemic countries, mainly the United States and Europe. Current anti-*T. cruzi* drugs have shown to be unable to block progression toward CCC [3].

After acute infection, parasitism is partially controlled by the immune response, and low-grade parasite persistence fuels the systemic production of inflammatory cytokines like IFN- γ and TNF- α , which is more intense in CCC than ASY patients [4–6]. CCC is characterized by a monocyte and T cell-rich myocarditis [7,8] with cardiomyocyte damage and hypertrophy, and prominent fibrosis; *T. cruzi* parasites are very scarce. IFN- γ producing Th1 cells accumulate in the myocardium of CCC patients [4,9,10] in response to locally produced chemokine ligands CXCL9 and CCL5 [11]. Accordingly, IFN- γ was found to be the most highly expressed cytokine mRNA in CCC myocardium using a 13-cytokine panel [12]. Both heart-crossreactive [13] and *T. cruzi*-specific T cells [14] have been found in CCC heart tissue, and both may play a role in the myocarditis of CCC. Together, evidence suggests that myocarditis and IFN γ signaling plays a major pathogenic role in CCC development and severity (reviewed in [2,15]), although downstream events leading to the heart disease phenotype are still obscure.

CCC has a worse prognosis than cardiomyopathies of non-inflammatory etiology, like ischemic or idiopathic dilated cardiomyopathy (DCM) [15]. Our group has shown that the myocardial gene expression profiles in CCC patients are profoundly different from those of both heart donors and DCM patients as assessed the “Cardiochip” cDNA microarray encoding ca. 11,000 expressed sequence tags (EST) cDNAs expressed in cardiovascular tissue [16]. Indeed, 15% of all genes specifically upregulated in CCC myocardium were found to be IFN- γ -inducible, indicating a strong IFN- γ transcriptional signature. This suggested that the increased aggressiveness of CCC could be related at least in part to activation of IFN- γ -dependent genes and pathways. Significantly, systemic overexpression of IFN γ in transgenic mice causes a TNF α -dependent inflammatory dilated cardiomyopathy [17,18]. Likewise, immunohistological signs of inflammation in suspected myocarditis of postviral etiology is associated with a poor prognosis [19], and sustained and long-term inflammation plays a role in worsening cardiac hypertrophy and chronic heart failure [20]. Mitochondrial dysfunction and oxidative stress have been associated with the pathogenesis of dilated cardiomyopathy and heart failure [21,22]. Indeed, our group has observed altered levels of 16S mitochondrial RNA as well as mRNA encoding mitochondrial proteins [16] and reduced levels of mitochondrial energy metabolism enzymes in that were specific to CCC [23]. Myocardial mitochondrial dysfunction and oxidative stress have been identified and explored in murine models of CCC (Reviewed in [24]). However, the determinants of expression of the majority of differentially expressed genes in CCC- as well as their roles in the key pathogenic roles of hypertrophy, fibrosis, arrhythmia and myocarditis—still remained mostly unknown.

MicroRNAs (miRNAs), short non-coding RNAs (18 to 24 nucleotides), are post-transcriptional regulators critically involved in a multitude of biological processes by modulating protein expression of up to 60% of the genes. miRNAs act by hybridizing with complementary sequences the 3' untranslated region (UTR) of mRNAs, exerting a downregulatory effect through direct degradation of the target mRNA and/or translational repression. MiRNAs also play a key role in multiple disorders, including cardiovascular disease [25]. Modulation of miRNA expression can profoundly alter disease phenotypes, and miRNA-based therapeutics has already entered clinical trials [26]. Our group recently studied the mRNA and miRNA transcriptome in the myocardium of mice acutely infected by *T. cruzi* [27], and we have previously shown that expression of muscle-enriched miRNAs (“myoMiRs”, including miR-1 and miR-133) is downmodulated in CCC myocardium, suggesting miRNA may control expression of pathogenetically relevant genes in CCC [28]. We raised the hypothesis that mRNA expression and pathways/processes may to be heavily influenced by miRNA expression. To comprehensively address this issue, we performed an integrative genome-wide analysis of the role of miRNA in global gene expression in CCC.

Methods

Ethics statement

The protocol was approved by the Institutional Review Board of the University of São Paulo, School of Medicine and written informed consent was obtained from the patients. All experimental methods comply with the Helsinki declaration.

Patients and sample collection

Human left ventricular free wall heart tissue was obtained from end-stage heart failure patients at the moment of heart transplantation. Patients with CCC presented positive *T. cruzi* serology and typical heart conduction abnormalities (right bundle branch block and/or left anterior division hemiblock) and had a histopathological diagnosis of myocarditis, fibrosis and

hypertrophy. Left ventricular free wall samples were also obtained from hearts of organ donors, which were not used for transplantation due to size mismatch with available recipients. (n = 4) All left ventricular free wall heart tissue samples were cleared from pericardium and fat and quickly frozen in liquid nitrogen and stored at -80°C .

RNA Extraction and RT-PCR

Myocardium samples (20–30 mg) were crushed with ceramic beads (CK14, diameter 1.4 mm) in 350 μl of RLT lysis buffer supplemented with 3.5 μl of β -mercapto-ethanol. Total RNA for mRNA expression profiling was extracted with the RNeasy Mini Kit (Qiagen, Courtaboeuf, France) adapted with Trizol. RNA quality and quantity was measured with a 2100 Bioanalyser. Total RNA (1 μg), with a RIN > 7, was reverse-transcribed with the high Capacity cDNA Reverse Transcription Kit (ThermoFisher Scientific, Saint Aubin, France).

Whole human genome expression analysis

Whole genome expression analysis was done on SurePrint G3 Human GeneExpression v1 8x60K arrays (Agilent Technologies, Les Ulis, France) following the manufacturer's protocol. Gene expression data were previously deposited in the GEO database ([GSE84796](#) and [GSE111544](#)). Microarray analyses and signal normalization were done with GeneSpring software (11.5.1), T test with adjustment for false discovery rate with the Benjamini-Hochberg method. Genes were considered differentially expressed if adjusted P values were <0.05 and absolute fold change >2.0. In order to validate the microarray results, quantitative real-time PCR, from 20ng of cDNA, was performed with the ABI 7900HT thermocycler and TaqMan Universal PCR Master Mix (Applied Biosystems, Life Technologies). The Student's T test was used to identify differentially expressed genes between CCC and controls by TaqMan RT-qPCR. Gene expression data used in this manuscript are the same ones that used in one of previous work on methylation analysis.

Principal component analysis, network and pathways analysis

Principal component analysis (PCA) analysis was performed using all differentially expressed genes and the variance expression of the number of standard deviations from mean overall samples. Canonical pathways analysis, networks analysis and Upstream regulator analysis, and classification of differentially genes belonging to pathways and biological processes were performed with Ingenuity Pathway Analysis (IPA, Qiagen Redwood City, CA, USA). We also classified genes in additional relevant pathobiological processes and pathways such as inflammation, IFN γ -modulated genes/Th1 response, extracellular matrix, fibrosis, hypertrophy, contractility of heart, hypertrophy, arrhythmia, oxidative stress/antioxidant response, mitochondria, and mitochondria-related genes using IPA Knowledge Base (IKB) gene lists, which were in some cases merged with other published gene lists. The IFN γ -dependent/Th1 response gene list was merged with published IFN γ -induced/repressed gene lists [29], and the oxidative stress gene list was merged with Nrf2-modulated genes [30]. The NF- κ B-modulated gene list was obtained from Yang et al. [31]. The mitochondrial gene list was a combination of all genes contained in the Mitochondrion Gene Ontology term and Mitocarta 2.0 [32]. Deconvolution of immune and cardiac cell types was performed by comparing the differentially expressed genes with the ARCHS4 tissue database in the EnrichR tool [33] (Adjusted P-value < 0.001). The network of cell types representing the genes shared by different tissues was constructed using the Cytoscape tool [34]. Prediction of differentially expressed miRNA-mRNA target relationships was performed with the IPA Knowledge Base. We selected high predicted or experimentally validated miRNA-target relationships.

Quantitative miRNA expression profiling

RNA and cDNA were obtained from human myocardial samples as previously described [35]. MiRNA profiling experiments were done for 754 miRNAs using pre-printed TLDA microfluidic cards (Human MicroRNA Card Set v3.0), according to the manufacturer's protocols and as described [35]. Raw TLDA data files were pre-processed with threshold and baseline corrections for each sample (automatic baseline and threshold set to 0.3) with assessment of each amplification plot on SDS 2.3 software (ThermoFisher). Cycle threshold (Ct) values from quantitative real time PCR data were imported, normalized and tested for statistical significance with the HTqPCR Bioconductor package [36]. Samples data quality and outliers removal were assessed with the arrayQualityMetrics Bioconductor package [37]. Distribution of the samples Ct values were normalized against the endogenous control RNU48-001006. Differentially expressed miRNAs (DEMs) were determined using a wrapper function from the Bioconductor package LIMMA [38]. Variance filtering was applied and, for each miRNA, up to two failed reads per group were accepted for partial coefficients calculation. Resulting p-values were submitted to false discovery rate adjustment according to the Benjamini-Hochberg method and the statistical significance threshold was defined as p-value ≤ 0.05 , with an absolute fold change cutoff ≥ 1.5 .

miRNA target gene interaction analysis

miRNA-target gene interaction analysis was done with performed with Ingenuity Pathway Analysis (IPA, Qiagen Redwood City, CA, USA). Analysis was done in three steps. First of all, for each DEM, we extracted from IPA database all the reported target genes (high predicted or experimentally observed). Then, among all these targets we kept only the ones that were differentially expressed (DEGs) in our gene expression analysis. Finally, we kept only the targets presenting an inverse pairing expression. In the analysis we included DEMs with an absolute fold change over 1.5 and DEGs with an absolute fold change over 2.0.

Results

Information on the subjects studied in this paper is available in [Table 1](#)

We found 1535 genes to be differentially expressed (DEG) between CCC and control myocardium, of which 1105 (72%) are upregulated, while 430 (28%) genes are downregulated in CCC ([S1 Table](#)). To validate the microarray results, we performed qPCR of 44 differentially expressed genes on the same samples (independent extractions) used for the microarray study, plus 16 new CCC samples. The confirmation rate was 86%, and only for 6 genes (ABRA, CDC42, ESRRA, GPD1, NFATC2, TGFBR2), the expression patterns were not confirmed. The gene specific qRT-PCR results, including fold change and p values, were previously described ([39], [S2 Table](#)).

A PCA analysis based on the all differentially expressed genes (DEGs) showed clustering of samples from each group in distinct areas of the plot ([Fig 1A](#)), confirming that CCC myocardial gene expression patterns were substantially different from controls. A heatmap based on DEGs confirmed the good clustering ([S1A Fig](#)). Likewise, differential clustering of CCC and control miRNAs also confirmed miRNA expression patterns are distinct in CCC and controls ([Fig 1B](#)). IPA canonical pathways analysis showed that the most enriched pathways are mainly immune-related, such as Th1 and Th2 T cells, dendritic cells/antigen presentation, leukocyte extravasation, NK and B cells; this is consistent with the high number of upregulated genes from the incoming inflammatory cells present in CCC but not in control heart tissue ([Fig 2A](#)). [S2 Table](#) depicts the DEGs belonging to all significant IPA canonical pathways. [Fig 2B](#) shows the number of genes in each pathobiological process relevant for the disease such as

Table 1. Characteristics of the human left ventricular free wall heart tissue samples used in this study.

| Project Number | Form | EF | Age | Sex | Transcriptome Analysis | qRT-PCR validation | MiRnome Analysis |
|----------------|---------|------|-----|-----|------------------------|--------------------|------------------|
| EBS | CCC | 0.12 | 32 | M | x | x | x |
| NSR | CCC | 0.15 | 49 | F | x | x | |
| MGS | CCC | 0.20 | 61 | F | | x | |
| BHAN | CCC | 0.20 | 15 | M | | x | |
| SCS | CCC | 0.17 | 59 | M | x | x | x |
| ECA | CCC | 0.19 | 32 | F | | x | |
| VTL | CCC | 0.19 | 41 | M | | x | |
| APA | CCC | 0.20 | 60 | F | x | x | |
| MCRS | CCC | 0.20 | 45 | F | x | x | x |
| MERS | CCC | 0.20 | 39 | F | | x | |
| MSS | CCC | 0.20 | 46 | F | | x | |
| GMS | CCC | 0.20 | 58 | M | | x | |
| ISM | CCC | 0.20 | 39 | M | | x | |
| OMG | CCC | 0.21 | 49 | M | | x | |
| MAP | CCC | 0.23 | 50 | F | x | x | |
| EPG | CCC | 0.23 | 41 | M | | x | |
| JRJ | CCC | 0.23 | 51 | M | | x | |
| LRJ | CCC | 0.25 | 66 | F | | x | |
| HBO | CCC | 0.25 | 36 | M | x | x | x |
| PMG | CCC | 0.29 | 57 | M | x | x | x |
| ABG | CCC | 0.30 | 64 | F | | x | |
| ZMC | CCC | 0.36 | 54 | F | x | x | x |
| JAB | CCC | 0.55 | 41 | M | | x | |
| AAF2 | CCC | 0.64 | 60 | M | x | x | x |
| JMS | CCC | 0.66 | 50 | M | | x | |
| LAL | CCC | 0.29 | 39 | M | | x | x |
| EMBT | control | | 25 | M | x | x | |
| LO | control | | 46 | M | x | x | |
| ESS | control | | 22 | M | x | x | x |
| ZFS | control | | x | M | x | x | x |
| FJR | control | | 28 | M | x | x | x |
| MBFM | control | | 17 | M | x | x | x |
| 3557 | control | | | | x | x | |

<https://doi.org/10.1371/journal.pntd.0008889.t001>

inflammation, IFN γ -modulated genes/Th1 response, extracellular matrix, fibrosis, contractility of heart, hypertrophy, arrhythmia, oxidative stress/antioxidant response, and mitochondria-related genes. The number of DEGs for each process is described on Fig 2B. S3 Table contains the complete list of DEGs belonging to each pathobiological process. As expected, inflammation and IFN γ -dependent/Th1 response processes show the highest number of DEGs (361 and 148, respectively), followed by fibrosis (82) and hypertrophy (53). Of interest, we found a significant number of DEGs belonging to mitochondria and oxidative stress functions/processes (42 and 35, respectively). Some DEGs are shared by several biological functions/processes (S4 and S5 Tables). IFN γ -dependent DEGs were found in all other 8 processes, ranging from 9% to 40% of genes in the other processes; those represented 104 inflammation, 33 fibrosis, 18 hypertrophy, 8 contractility and 7 mitochondrial genes.

Among these processes inflammation may be specific to CCC as shown in previous gene expression studies [11,12]. For the other processes, they have been also described in dilated

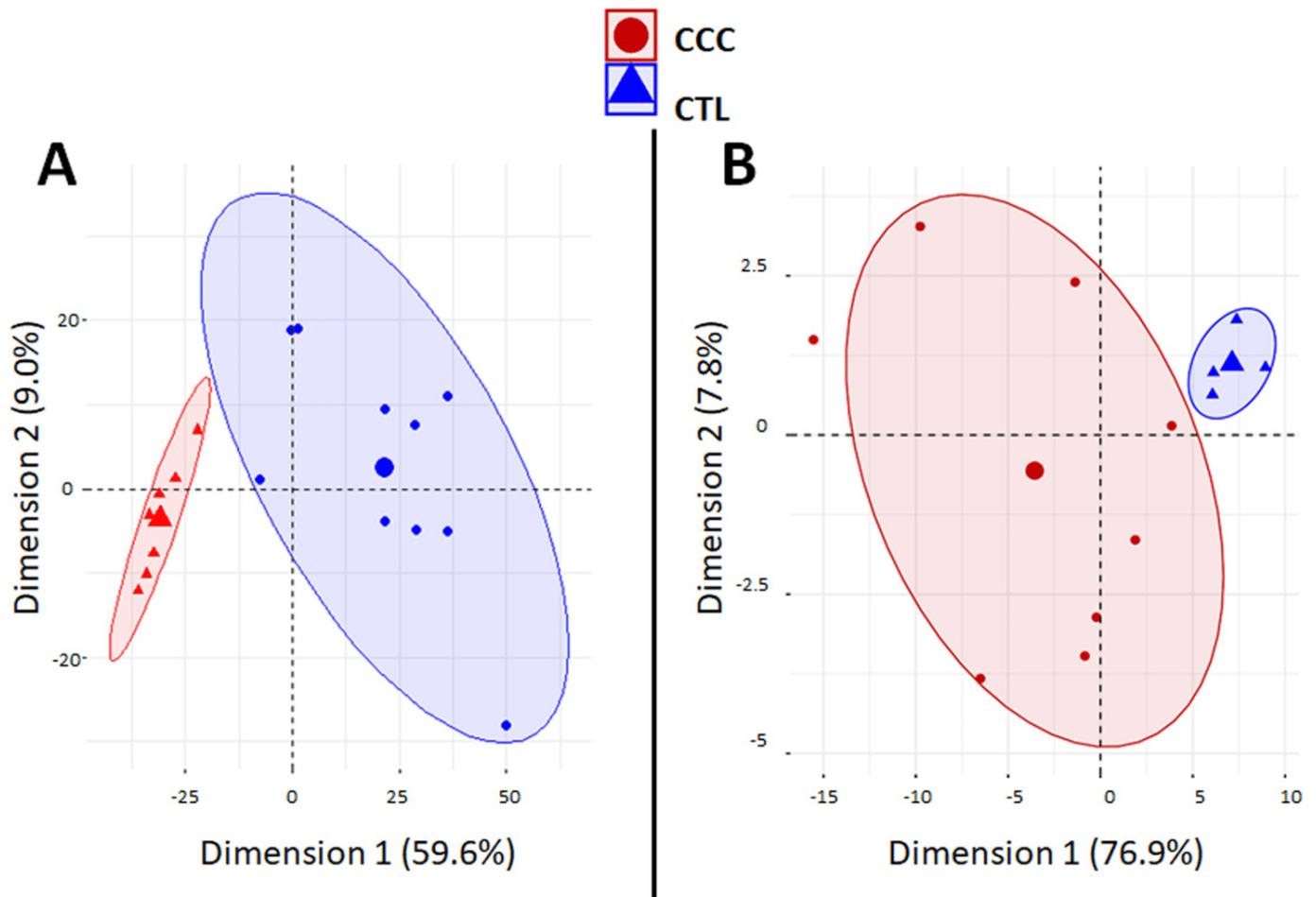


Fig 1. Principal component analysis (PCA) plots. Principal component analysis (PCA) plot of samples was performed based (A) on 1535 differentially expressed genes (DEGs) between CCC and controls. The two main principal components have the largest possible variance (68.8%). The 1535 DEGs had an equal contribution to the first component (ranging from 6.0E-3% to 0.1%). For the second component, 25 DEGs had a contribution over 0.25% (GAB3, WBSCR27, LOC100130930, C1orf35, ISLR2, SLC25A34, NOTCH2, TSPAN32, ATP1A1OS, C11orf65, ZNF214, APCDD1, C1QTNF6, RANBP17, MNS1, APBB3, ANGPTL1, BEND6, LTB, MMP9, ITGB2, PIK3R1, NOTCH2NL, TRMT5 and XLOC_005730); (B) or on 80 differentially expressed miRNAs (DEMs) between CCC and controls. The two main components explain 84.7% of the variance. For the first component, the 80 DEMS had an equal contribution (ranging from 0.24% to 1.74%). For the second component, even if all the DEMs contribute, six of the DEMs have a main contribution (hsa-miR-155: 12.8%; hsa-miR-146a: 9.0%; hsa-miR-302d: 8.8%; hsa-miR-378: 8.1%; hsa-miR-486: 7.0%; and hsa-miR-221: 5.2%).

<https://doi.org/10.1371/journal.pntd.0008889.g001>

cardiomyopathies of other etiologies. Upstream regulator analysis performed by IPA examines how many targets of each given transcriptional regulator are present in the DEGs—as well as the direction of change—based on the literature and IPA knowledge base; putative regulators are ranked according to overlap with expected targets and directionality (z -score). It indicated that $IFN\gamma$ is the top upstream regulator, followed by other cytokines like $TNF\alpha$, IL-18 and EBI3/IL27R β chain, the chemokines CCL5 and CXCL10, the transcription factors NF- κ B and Ap1, and the PI3K enzyme (Table 2). S6 Table shows the 27 cytokines and chemokines upregulated in CCC heart tissue. Significantly, the 7 most upregulated among them were chemokines, including chemokine ligands of CCR5 (CCL5, CCL4) and CXCR3 CXCL9 and CXCL10). Multiple cytokines and chemokines that were top upstream regulators like $IFN\gamma$, CCL5, CXCL10, IL-18, IL-7, EBI3/IL-27b and IL-4 were found to be upregulated to different degrees in CCC myocardium. Deconvolution of immune cell type profiles in CCC

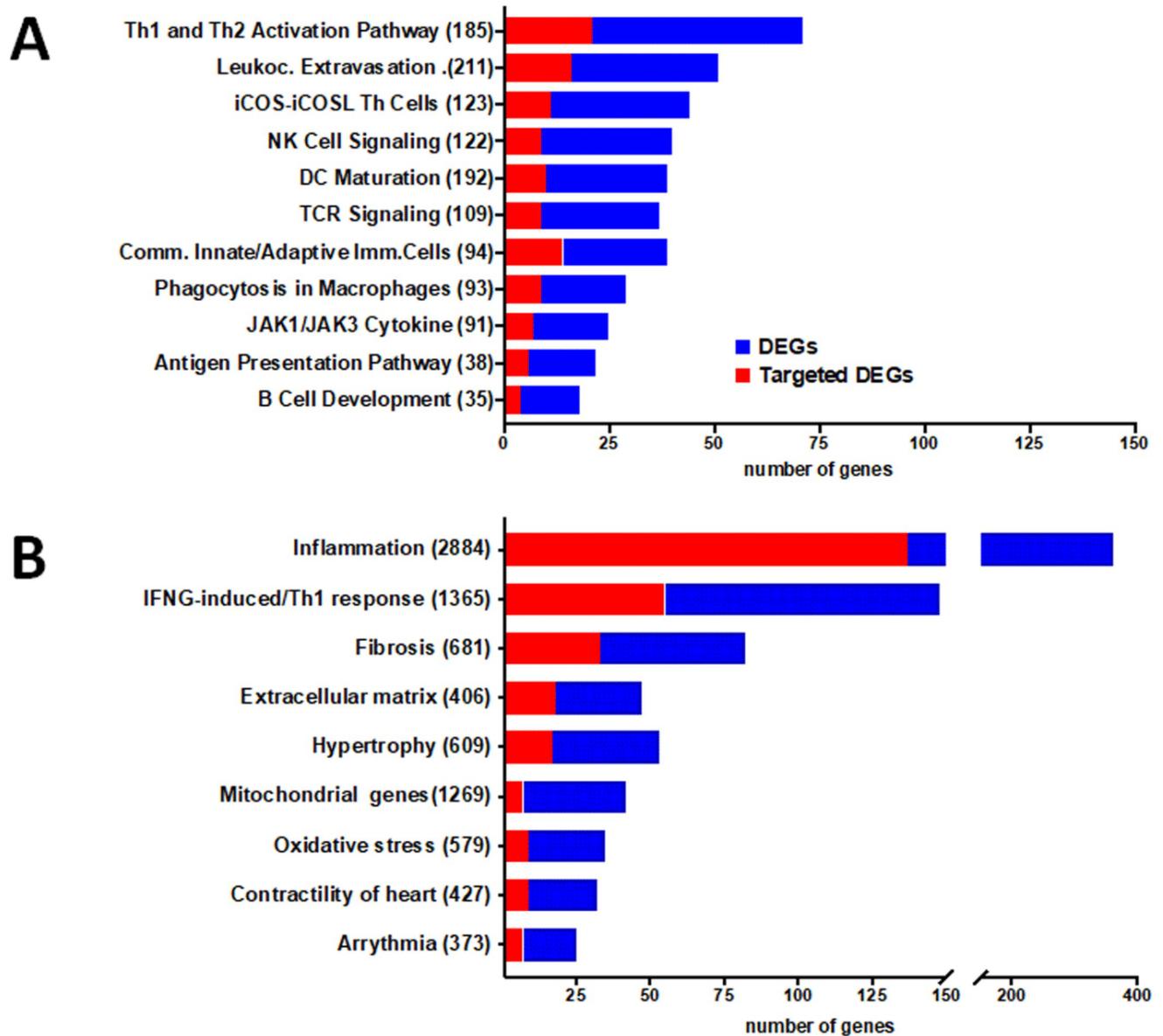


Fig 2. DEGs and DEM-targeted DEGs present in relevant canonical pathways and pathophysiological processes. The stacked bar chart displays the number of DEG (blue) and DEM-targeted DEGs (red) present in each pathway. **A.** Ingenuity Pathway Analysis (IPA) canonical pathways representative of the most significantly enriched in the heart of CCC patients. **B.** DEG and DEM-targeted DEGs in specific biological processes relevant for the CCC pathogenesis. The numerical value in the parentheses in front of each pathway name represents the total number of genes in that pathway/process.

<https://doi.org/10.1371/journal.pntd.0008889.g002>

myocardium revealed an enrichment of gene expression signatures of CD4+ T cells, NK cells, B cells/plasma cells, dendritic cells, plasmacytoid dendritic cells, regulatory T cells and granulocytes (red; Fig 3). This indicates that these cell types infiltrate the myocardium of CCC patients. Conversely, genes down-regulated in CCC myocardium when compared to controls were enriched with signatures of cardiac muscle cells (blue; Fig 3). This result is most likely a consequence of reduced representation of cardiac mRNAs in CCC myocardium that was replaced by inflammatory cells.

Table 2. Upstream regulator analysis in CCC myocardium.

| Upstream Regulator | Molecule | Activation z-score | p-value of overlap |
|-----------------------------------|----------|--------------------|--------------------|
| IFNG | cytokine | 7.891 | 2.38E-24 |
| TNF | cytokine | 6.494 | 2.19E-12 |
| IL18 | cytokine | 4.583 | 8.95E-14 |
| NFkB (complex) | complex | 4.465 | 5.54E-06 |
| CD40LG | cytokine | 4.287 | 4.10E-15 |
| TCR | complex | 3.832 | 1.80E-26 |
| BCR (complex) | complex | 3.771 | 1.27E-15 |
| IL7 | cytokine | 3.466 | 6.32E-18 |
| TET2 | enzyme | 3.357 | 4.04E-03 |
| IL4 | cytokine | 3.082 | 7.43E-23 |
| EBI3 (IL27R β chain IL27RB) | cytokine | 2.785 | 2.13E-05 |
| Fcer1 | complex | 2.72 | 3.35E-05 |
| CXCL10 | cytokine | 2.581 | 3.15E-04 |
| Ap1 | complex | 2.454 | 5.70E-03 |
| CCL5 | cytokine | 2.432 | 5.92E-03 |
| U1 snRNP | complex | 2.384 | 1.24E-06 |
| TNFSF14 (LIGHT) | cytokine | 2.224 | 9.77E-03 |
| PI3K (complex) | complex | 2.007 | 3.62E-04 |
| Collagen type I | complex | 2.000 | 3.96E-01 |

<https://doi.org/10.1371/journal.pntd.0008889.t002>

Regarding microRNA analysis, it was performed on 8 CCC and 4 control myocardial samples contained in the mRNA transcriptome experiment (10 CCC and 7 control myocardial samples). 754 human miRNAs were screened on the heart samples and among them, 210 miRNAs were detected in every sample; these were quantified in each tissue sample. **S7 Table** shows the expression levels and statistical significance of the 210 expressed miRNAs. Based on their expression values, we have found that 80 out of 210 miRNAs were differentially expressed (DEMs) (absolute FC ≥ 1.5 , $p < 0.05$ without correction). A PCA analysis based on the all differentially expressed miRNAs (**Fig 1B**) and a heatmap (**S1B Fig**) confirmed that miRNA patterns were substantially different from controls. After correction for multiple testing, only miR-146a-000468 ($p = 6.9E-03$) and miR-155-002623 ($p = 6.9E-03$) remain significantly altered. However, the list of the 80 miRNAs obtained without correction for multiple testing seems to be relevant as it contains miR-1 ($p = 5.0E-03$), miR-133a ($p = 1.4E-02$) and miR-133b ($p = 1.3E-2$) that we previously observed as under-expressed in CCC samples as compared to controls [28]. MiR-208a, which was also previously described to be under expressed in CCC samples in the same study [28], is borderline in the present study ($p = 5.5E-02$).

As the controls were younger than CCC we performed some PCA analyses restricted on CCC including DEG information (**S2A Fig**) or DEM information (**S2B Fig**) then we overlaid the age of the patients. No obvious correlation was detected. We performed some spearman correlation tests 9 DEGs were age correlated and none of the DEM were age correlated (**S8 Table**). So, the age may not act as a confounding factor. Similarly, we did a PCA analysis on cases and controls taking into account DEG information or DEM information and the sex of the patients. No specific clustering was detected (**S3A and S3B Fig**). For each DEG and DEM we made Student's t tests between the male and female patients and no association were detected.

In order to identify putative miRNA-target gene interactions among DEMs and DEGs, we performed inverse expression pairing of DEMs (80) and DEGs (1535). A total of 571 miRNA-

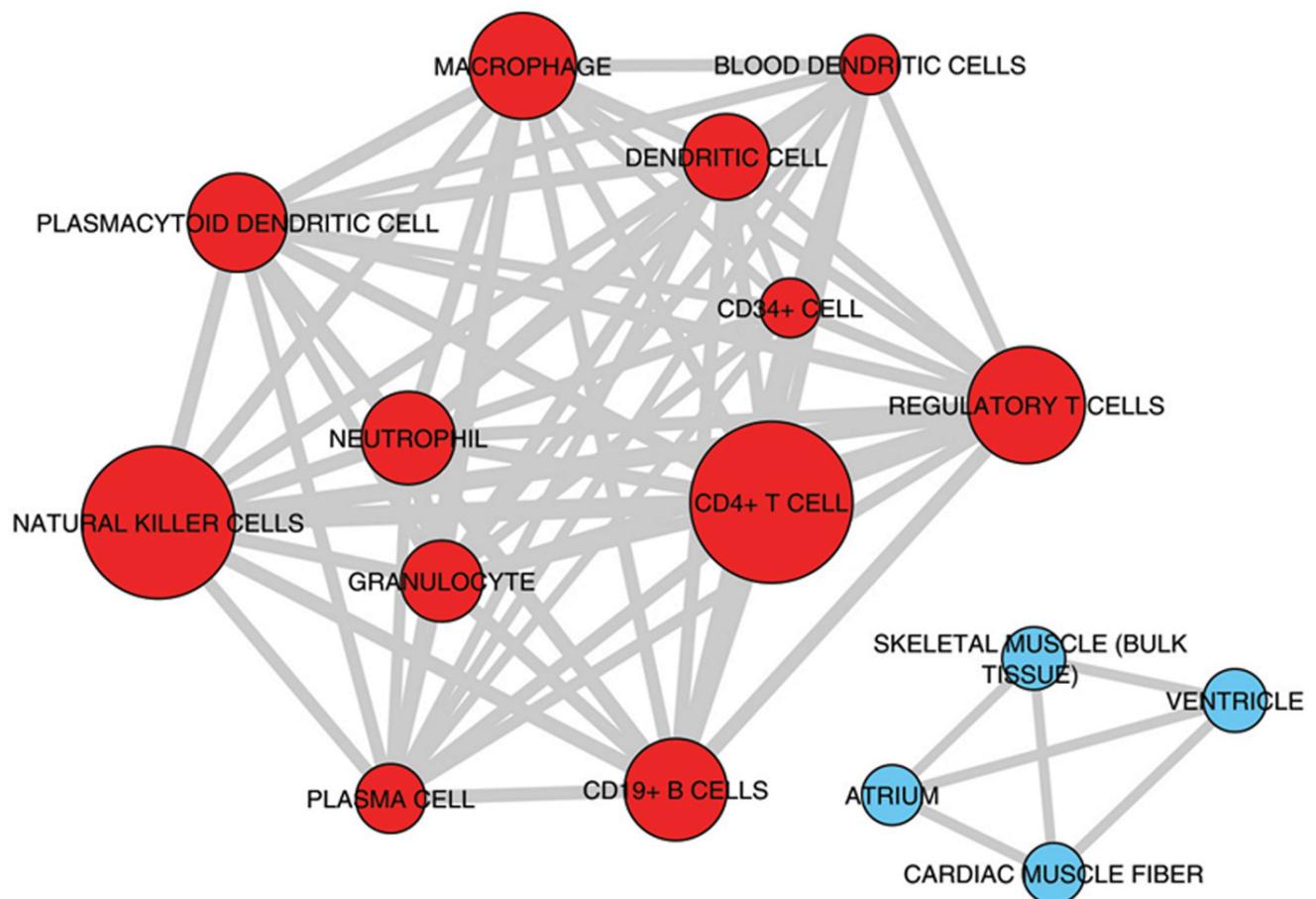


Fig 3. Enrichment analysis of cell subset and tissue signatures. Signatures of different tissues and cell types from ARCHS4 tissue database enriched with up-regulated (red nodes) or down-regulated (blue nodes) genes compared to controls (Adjusted P-value < 0.001). The width of edges (connecting lines) is proportional to the number of genes shared by two signatures. The size of nodes is proportional to the $-\log_{10}$ Adjusted P-value.

<https://doi.org/10.1371/journal.pntd.0008889.g003>

mRNA interactions involving 67 DEMs and 396 DEGs were found by IPA. **S9 Table** depicts all observed DEM-DEG interactions. **Fig 2A** shows the number of DEGs in the most important canonical pathways and the fraction that is targeted by DEMs. The proportion of DEGs targeted by DEMs in each depicted canonical pathway varies from 24% to 62%. A similar analysis done on DEM-DEG interactions in the 9 key pathobiological processes, which also indicated that a substantial proportion of DEGs (16.7–40.8%) are targeted by DEMs. Apart from the inflammation (38%) and $\text{IFN}\gamma$ -induced genes (37.2%), pathobiological processes with the highest number of DEM-targeted DEGs are fibrosis (40.2%), extracellular matrix (38.2%) and hypertrophy (32.1%) processes (**Fig 2B**). **S10 Table** shows all DEM-DEG interactions classified according to biological process. In order to validate the specificity of the DEM-DEG targeting, we simulated the targeting of the DEGs with 80 miRNAs that were not differentially expressed in CCC myocardium (non-DEMs) as compared to 80 DEMs using the IPA miRNA target filter function, again focusing only high predicted and experimentally observed targets. Comparing the number of target DEGs in the top 6 IPA canonical pathways that were shared by DEMs and non-DEMs, we found a 50% higher number of DEG targets from DEM than DEG targets of the simulated non-DEMs ($p < 0.00001$, chi-square). This suggests that the pairing of DEG targets with DEM was not random.

We found that 5 miRNAs (hsa-miR-125b-5p, hsa-miR-15a-5p, hsa-miR-296-5p, hsa-miR-29c-3p and hsa-miR-103a-3p) each regulate more than twenty DEGs; moreover, each of them affects at least 6 of the 9 biological functions and processes analyzed (S11 Table). Moreover, several of these “master” miRNAs targeted multiple genes belonging to a given process at the same time, suggesting a synergistic action. A network built with DEM-DEG targets around the important pathobiological processes, myocarditis, fibrosis, hypertrophy and arrhythmia disclosed a strong focus on fibrosis, and several miRNAs and targets participated in various processes (summarized on Fig 4). We validated the expression of 38 genes belonging to the four pathobiological processes with real time RTqPCR in a larger set of CCC samples.

Discussion

To assess the role of miRNAs in regulating gene expression in CCC myocardium, we performed an integrative genome-wide analysis of miRNAs and mRNA expression in CCC myocardium samples and performed network and pathways analysis. We identified 1535 differentially expressed genes (DEGs) and 80 differentially expressed miRNAs (DEMs). We found that both miRNAs and mRNA expression profiles discriminated CCC from control samples. Pathways analysis disclosed an enrichment in inflammation, Th1/IFN- γ -inducible genes, genes belonging to fibrosis, hypertrophy, and mitochondrial/oxidative stress/antioxidant response. Our results corroborated that IFN- γ is the key cytokine modulating transcriptional changes in CCC myocardium and affecting all other studied pathobiological processes, and cell type deconvolution indicated the presence of novel immune cell types that had not yet been disclosed by immunohistochemistry. Our data also suggest that a significant number of differentially expressed microRNAs target differentially expressed genes; moreover, a few microRNAs may potentially regulate simultaneously multiple genes in key pathways and pathogenetically relevant processes. Our paper is the first to indicate that miRNAs may play a role in promoting major transcriptome changes in human inflammatory cardiomyopathy.

Our results pointed out IFN- γ is the top gene expression regulator with ca. 10% of DEGs being modulatable by it, in all pathobiological processes. Indeed, several studies have shown a negative impact of IFN- γ on the myocardium, leading to reduced contractility, release of chemokines and increased production of atrial natriuretic factor [40–42]. IFN- γ -induced cardiac fibrosis with increased fibroblast proliferation, production of hyaluronan and metalloproteinases 2 and 9 has also been demonstrated [43–46]. The role of IFN- γ , TNF- α and NF- κ B as top upregulators are also in line with data in genetically modified murine models. Mice transgenic to IFN- γ developed a TNF- α -dependent inflammatory dilated cardiomyopathy with fibrosis and heart failure [18], and a very similar phenotype was developed by mice constitutively expressing active IKK2 [47]. Mechanistically, IFN- γ induces TNF- α and potentiates TNF- α -mediated NF- κ B signaling and upregulation of NOS2 [48,49], leading to cardiomyocyte contractile dysfunction and apoptosis [Sun, 1998 #110]. This is mediated at least in part by NADPH- and NOS2-dependent production of reactive oxygen and nitrogen species (ROS and RNS, respectively), with oxidative and nitrosative stress [40,50]. IFN- γ -induced RNS leads to inhibition of mitochondrial oxidative metabolism [51] and ATP depletion in cardiomyocytes [52] with ensuing mitochondrial dysfunction. Of interest, 169 DEGs, or ca 10% of DEGs are potentially modulated by NF- κ B in CCC myocardium. Our data point towards IFN- γ and NF- κ B-mediated signaling as a major player in Chagas cardiomyopathy; we believe they may have a central role in orchestrating the molecular processes that contribute to heart failure. It is noteworthy that IFN- γ may also act through modulation of miRNA expression in CCC myocardium. IFN- γ down-regulates 5 miRNAs (27b, 92a, 99a, 99b, 101) [53,54] which found to be downregulated DEMs in CCC myocardium.

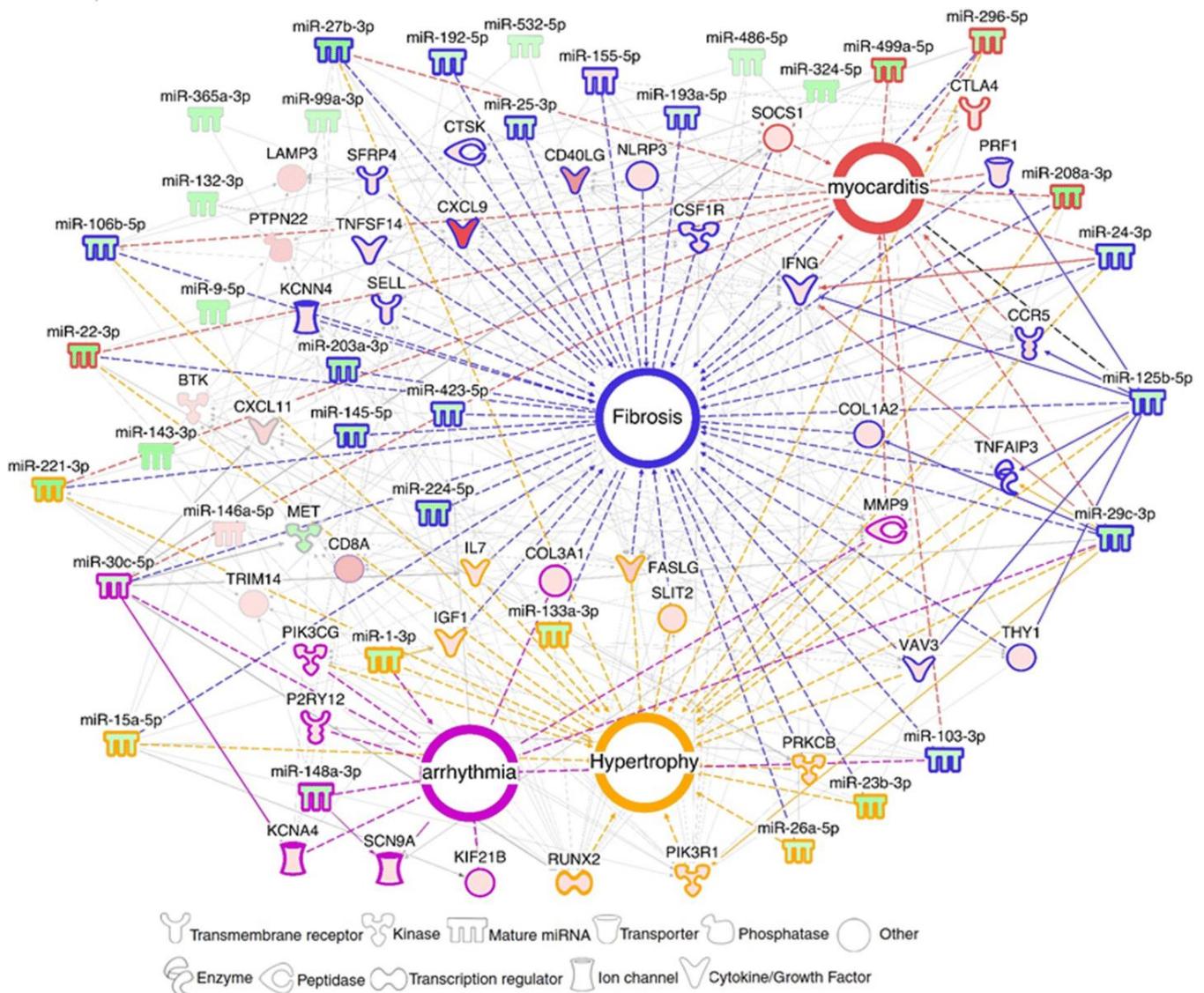


Fig 4. DEM-DEG network related to the main Chagas disease pathobiological features. Networks with DEMs and DEGs related to myocarditis, fibrosis, arrhythmia and hypertrophy were built using IPA software. Each built network contains molecules represented as nodes, and the biological relationship between two nodes is represented as an edge (line). All edges (connecting lines) are interactions supported by at least one reference from the literature or from canonical information stored in the IPA Ingenuity Knowledge Base (IKB). Full lines are direct interactions, dotted lines are indirect interactions. Upregulated microRNAs and mRNAs are colored in hues of red, and downregulated molecules are colored in hues of green according to the intensity of expression. Each node shape represents one type of molecule.

<https://doi.org/10.1371/journal.pntd.0008889.g004>

Previous immunohistochemistry studies on CCC heart tissue have identified a few cell types in the inflammatory infiltrate. Most cells were CD68+ macrophages, CD4 and CD8+ T cells, with fewer B cells, NK cells, and TGF-beta-expressing cells [7,8,10]. In addition to these cell types, the deconvolution of immune cell type transcriptional profiles shown here provided evidence for the presence of plasmacytoid dendritic cells, regulatory cells, plasma cells and granulocytes. Many of the genes in the NK signature are shared with CD8+ cytotoxic T cells which are abundant in CCC heart tissue [8]; this has previously been observed in transcriptome profiling of peripheral blood from CCC patients [27], and we believe our finding in

myocardium represents an NK/CD8⁺ T cell cytotoxicity signature. Regarding regulatory T cells, we had previously found a low expression of CTLA-4 mRNA in CCC heart tissue, suggesting a small component of CTLA-4⁺ T regs [12].

A significant proportion of the DEGs in the pathways and processes we studied (15–62%) were targeted by differentially expressed miRNA (DEM). We found that some DEMs had unusually high numbers of target DEGs. Five downmodulated DEMs (hsa-miR-15a-5p, hsa-miR-29c-3p, hsa-miR-103a-3p, hsa-miR-125b-5p and hsa-miR-296-5p) each targeted at least 20 DEGs involved in 6 or more studied pathobiological processes. Their functions in the context of heart disease and inflammation are described below. Downmodulated miR-15a may target 23 upregulated DEGs involved in 5 out of the 8 processes described above. Previous data indicate that the miR-15 family is involved in TGF β -pathway inhibition [55] while miR-15 antagonists induce fibrosis and hypertrophy [56], which are important features of Chagas disease. Our study also revealed down regulation of hsa-miR-29c-3p, associated with 21 upregulated target DEGs involved in all processes described above. The miR-29 family members target several genes related to extracellular matrix and fibrosis [57–59], and miR29b was shown to be inhibited by the myocardial infarction associated transcript (MIAT) [29], a long non-coding RNA overexpressed in CCC patients' myocardium [60].

Downmodulated miR-103 targets 20 upregulated DEGs, involved in 6 out of the 8 processes previously described. MiR-103 is overexpressed in the myocardium of heart failure patients [61], and attenuates cardiomyocyte hypertrophy by a mechanism that partially relies on reducing cardiac autophagy [62]. Down modulated miR-125b targets 23 upregulated DEGs in 6 out of the 8 processes described above. The miR-125 family members negatively regulate the expression of TNF- α , reducing ischemia/reperfusion damage [63] also reducing chemokine RANTES (CCL5), which is highly expressed in our CCC samples. Lastly, down modulated miR-296 targets 22 upregulated DEGs. MiR-296 is linked to fibrosis, and is similarly downregulated in hypertensive patients [64].

Only two miRNAs (miR-155-5p and miR-146a-5p) were found to be upregulated in CCC patients. miR-155 has been shown to increase the global Nrf2 transcriptional response by targeting translation of the transcriptional regulator BACH1 [65], indicating that this miRNA may have an impact on oxidative stress in CCC hearts and its upregulation may have occurred as a compensatory mechanism to intense oxidative stress. This hypothesis will require biological validations. The Nrf2 pathway and HMOX1 have been reported to play a role in "tissue tolerance"—the ability of resist pathogen, inflammation, or oxidative stress-mediated damage during infection or inflammation [66,67] and we have found that HMOX2, a homologous Nrf2-induced gene involved in the antioxidant response, is down regulated in CCC myocardium. miR-146a is expressed in multiple cardiac cell types [68] and was found to be increased upon induction of cardiotoxicity and to inhibit proteins involved in heart regeneration (ErbB-2 and -4) [69,70], suggesting that overexpression of this miRNA may have a negative impact on cardiac function. On the other hand, miR-146a was shown to be induced by NF- κ B and to create an anti-inflammatory feedback loop by inhibiting NF- κ B-induced proinflammatory cytokine production and inflammatory cell migration into the myocardium [71,72] promoting Treg suppressor function [73], suggesting that this miRNA also presents cardioprotective effects. However, we have shown here and elsewhere [11,12] that Th1 inflammatory cytokines and chemokines are highly expressed in CCC heart tissue and very few Treg are detectable in CCC myocardium [10–12]. Like miR-155, it is possible that it is upregulated as a failed attempt to modulate inflammation.

Our study was performed in whole heart tissue, containing several cell types, including cardiomyocytes, fibroblasts, endothelial and infiltrating inflammatory cells. We must thus keep in mind that results reflect the composite of mRNA and microRNA content of each cell type with

its respective contribution. Most of the RNA will come from cardiomyocytes, but inflammatory cell RNA will readily show up, since control tissue is free from inflammatory infiltrates, showing at most passenger leukocytes that are much less numerous. At any event, our results suggest that, by targeting multiple genes in relevant pathogenic disease pathways and processes, miRNAs can exert a combined regulatory effect that may be stronger than the effect of a single DEM-DEG interaction. In addition, we found a small number of key "high-ranking" differentially expressed miRNAs—those with the highest number of targets, overlapping with those with multiple targets involved in several pathological processes. Our data identified specific molecular features in key pathogenic processes. Further investigation and validation of the more important miRNA-mRNA interactions involved in fibrosis, oxidative stress, and mitochondrial processes may reveal important insights into the pathogenesis of CCC and may translate in the identification of novel therapeutic targets. Our findings may have a bearing on myocarditis and inflammatory cardiomyopathy of distinct etiologies as well as to IFN- γ mediated age-related myocardial inflammation and functional decline [74] as recently described.

Supporting information

S1 Table. List of differentially expressed genes on human heart biopsies from end stage patients or organ donors.

(PDF)

S2 Table. Description of the canonical pathways containing the DEGs in CCC myocardium.

(PDF)

S3 Table. DEGs associated to biological functions and processes in CCC myocardium.

(PDF)

S4 Table. Number of DEGs shared between several pathobiological functions or processes in CCC myocardium.

(PDF)

S5 Table. DEGs shared between several pathobiological functions or processes in CCC myocardium.

(PDF)

S6 Table. Cytokines and chemokines differentially expressed in CCC heart tissue.

(PDF)

S7 Table. DEMs in CCC myocardium.

(PDF)

S8 Table. Correlation between DEG/DEM and the age of the patients.

(PDF)

S9 Table. DEM-DEG interactions in CCC myocardium.

(PDF)

S10 Table. DEM-DEG interaction in each pathobiological function or process in CCC myocardium.

(PDF)

S11 Table. Five DEMs that control a large number of DEGS and processes in CCC myocardium.

(PDF)

S1 Fig. Unsupervised hierarchical clustering done on patients with severe chronic Chagas disease cardiomyopathy (CCC) and controls. **A.** Unsupervised hierarchical clustering based on the 1535 DEGs. **B.** Unsupervised hierarchical clustering based on the 80 DEMs. (TIF)

S2 Fig. Principal component analysis (PCA) plots taking into account the age of the cases. Principal component analysis (PCA) plot of samples was performed based **A.** on 1535 differentially expressed genes (DEGs) between CCC and controls. **B.** on 80 differentially expressed miRNAs (DEMs) between CCC and controls. Each plot was generated only on cases and the age of the patients was overlaid. (TIF)

S3 Fig. Principal component analysis (PCA) plots taking into account the sex of the cases and controls. Principal component analysis (PCA) plot of samples was performed based **A.** on 1535 differentially expressed genes (DEGs) between CCC and controls. **B.** on 80 differentially expressed miRNAs (DEMs) between CCC and controls. On each plot cases and controls are indicated according to their sex. (TIF)

Acknowledgments

We thank all the participants and their relatives.

Author Contributions

Conceptualization: Jorge Kalil, Edecio Cunha-Neto, Christophe Chevillard.

Data curation: Sandrine Cabantous, Amanda Farage Frade, Priscila Camillo Teixeira, Darlan da Silva Cândido.

Formal analysis: Laurie Laugier, Ludmila Rodrigues Pinto Ferreira, Frederico Moraes Ferreira, Sandrine Cabantous, Amanda Farage Frade, Priscila Camillo Teixeira, Darlan da Silva Cândido, Edecio Cunha-Neto, Christophe Chevillard.

Funding acquisition: Jorge Kalil, Edecio Cunha-Neto, Christophe Chevillard.

Investigation: Laurie Laugier, Ludmila Rodrigues Pinto Ferreira, Sandrine Cabantous, Amanda Farage Frade, Priscila Camillo Teixeira, Darlan da Silva Cândido.

Methodology: Laurie Laugier, Ludmila Rodrigues Pinto Ferreira, Sandrine Cabantous, Amanda Farage Frade, Priscila Camillo Teixeira, Darlan da Silva Cândido.

Project administration: Edecio Cunha-Neto, Christophe Chevillard.

Resources: Ronaldo Honorato Barros Santos, Edimar A. Bocchi, Fernando Bacal, Edecio Cunha-Neto.

Software: Frederico Moraes Ferreira, Pauline Brochet, Vanessa Escolano Maso, Helder I. Nakaya.

Supervision: Jorge Kalil, Edecio Cunha-Neto, Christophe Chevillard.

Validation: Pauline Brochet, Vanessa Escolano Maso, Helder I. Nakaya, Edecio Cunha-Neto, Christophe Chevillard.

Visualization: Vanessa Escolano Maso, Helder I. Nakaya, Edecio Cunha-Neto, Christophe Chevillard.

Writing – original draft: Edecio Cunha-Neto, Christophe Chevillard.

Writing – review & editing: Joao Paulo Nunes, Rafael Almeida Ribeiro, Jorge Kalil, Edecio Cunha-Neto, Christophe Chevillard.

References

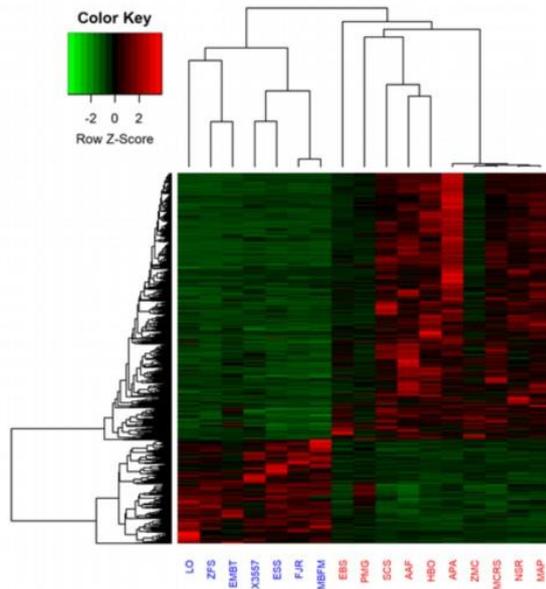
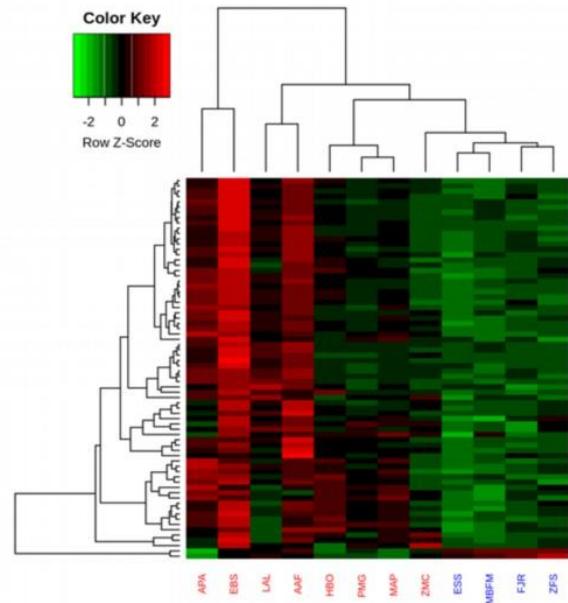
1. Bocchi EA, Bestetti RB, Scanavacca MI, Cunha Neto E, Issa VS Chronic Chagas Heart Disease Management: From Etiology to Cardiomyopathy Treatment. *J Am Coll Cardiol* 2017; 70: 1510–1524. <https://doi.org/10.1016/j.jacc.2017.08.004> PMID: 28911515
2. Chevillard C, Nunes JPS, Frade AF, Almeida RR, Pandey RP, Nascimento MS et al. Disease Tolerance and Pathogen Resistance Genes May Underlie *Trypanosoma cruzi* Persistence and Differential Progression to Chagas Disease Cardiomyopathy. *Front Immunol* 2018; 9: 2791. <https://doi.org/10.3389/fimmu.2018.02791> PMID: 30559742
3. Morillo CA, Marin-Neto JA, Avezum A, Sosa-Estani S, Rassi A, Rosas F et al. Randomized Trial of Benznidazole for Chronic Chagas' Cardiomyopathy. *N Engl J Med* 2015; 373: 1295–1306. <https://doi.org/10.1056/NEJMoa1507574> PMID: 26323937
4. Abel LC, Rizzo LV, Ianni B, Albuquerque F, Bacal F, Carrara D et al. Chronic Chagas' disease cardiomyopathy patients display an increased IFN-gamma response to *Trypanosoma cruzi* infection. *J Autoimmun* 2001; 17: 99–107. <https://doi.org/10.1006/jaut.2001.0523> PMID: 11488642
5. Gomes JA, Bahia-Oliveira LM, Rocha MO, Busek SC, Teixeira MM, Silva JS et al. Type 1 chemokine receptor expression in Chagas' disease correlates with morbidity in cardiac patients. *Infect Immun* 2005; 73: 7960–7966. <https://doi.org/10.1128/IAI.73.12.7960-7966.2005> PMID: 16299288
6. Sousa GR, Gomes JA, Fares RC, Damasio MP, Chaves AT, Ferreira KS et al. Plasma cytokine expression is associated with cardiac morbidity in chagas disease. *PLoS One* 2014; 9: e87082. <https://doi.org/10.1371/journal.pone.0087082> PMID: 24603474
7. Fonseca SG, Reis MM, Coelho V, Nogueira LG, Monteiro SM, Mairena EC et al. Locally produced survival cytokines IL-15 and IL-7 may be associated to the predominance of CD8+ T cells at heart lesions of human chronic Chagas disease cardiomyopathy. *Scand J Immunol* 2007; 66: 362–371. <https://doi.org/10.1111/j.1365-3083.2007.01987.x> PMID: 17635814
8. Higuchi Mde L, Gutierrez PS, Aiello VD, Palomino S, Bocchi E, Kalil J et al. Immunohistochemical characterization of infiltrating cells in human chronic chagasic myocarditis: comparison with myocardial rejection process. *Virchows Arch A Pathol Anat Histopathol* 1993; 423: 157–160. <https://doi.org/10.1007/BF01614765> PMID: 7901937
9. Reis MM, Higuchi Mde L, Benvenuti LA, Aiello VD, Gutierrez PS, Bellotti G et al. An in situ quantitative immunohistochemical study of cytokines and IL-2R+ in chronic human chagasic myocarditis: correlation with the presence of myocardial *Trypanosoma cruzi* antigens. *Clin Immunol Immunopathol* 1997; 83: 165–172. <https://doi.org/10.1006/clin.1997.4335> PMID: 9143377
10. Rocha Rodrigues DB, dos Reis MA, Romano A, Pereira SA, Teixeira Vde P, Totes S Jr et al. In situ expression of regulatory cytokines by heart inflammatory cells in Chagas' disease patients with heart failure. *Clin Dev Immunol* 2012; 2012: 361730. <https://doi.org/10.1155/2012/361730> PMID: 22811738
11. Nogueira LG, Santos RH, Ianni BM, Fiorelli AI, Mairena EC, Benvenuti LA et al. Myocardial chemokine expression and intensity of myocarditis in Chagas cardiomyopathy are controlled by polymorphisms in CXCL9 and CXCL10. *PLoS Negl Trop Dis* 2012; 6: e1867. <https://doi.org/10.1371/journal.pntd.0001867> PMID: 23150742
12. Nogueira LG, Santos RH, Fiorelli AI, Mairena EC, Benvenuti LA, Bocchi EA et al. Myocardial gene expression of T-bet, GATA-3, Ror-gamma, FoxP3, and hallmark cytokines in chronic Chagas disease cardiomyopathy: an essentially unopposed TH1-type response. *Mediators Inflamm* 2014; 2014: 914326. <https://doi.org/10.1155/2014/914326> PMID: 25152568
13. Cunha-Neto E, Coelho V, Guilherme L, Fiorelli A, Stolf N, Kalil J Autoimmunity in Chagas' disease. Identification of cardiac myosin-B13 *Trypanosoma cruzi* protein crossreactive T cell clones in heart lesions of a chronic Chagas' cardiomyopathy patient. *J Clin Invest* 1996; 98: 1709–1712. <https://doi.org/10.1172/JCI118969> PMID: 8878420
14. Fonseca SG, Moins-Teisserenc H, Clave E, Ianni B, Nunes VL, Mady C et al. Identification of multiple HLA-A*0201-restricted cruzipain and FL-160 CD8+ epitopes recognized by T cells from chronically *Trypanosoma cruzi*-infected patients. *Microbes Infect* 2005; 7: 688–697. <https://doi.org/10.1016/j.micinf.2005.01.001> PMID: 15848276

15. Cunha-Neto E, Chevillard C Chagas disease cardiomyopathy: immunopathology and genetics. *Mediators Inflamm* 2014; 2014: 683230. <https://doi.org/10.1155/2014/683230> PMID: 25210230
16. Cunha-Neto E, Dzau VJ, Allen PD, Stamatou D, Benvenuti L, Higuchi ML et al. (2005) Cardiac gene expression profiling provides evidence for cytokinopathy as a molecular mechanism in Chagas' disease cardiomyopathy. *Am J Pathol* 2005; 167: 305–313. [https://doi.org/10.1016/S0002-9440\(10\)62976-8](https://doi.org/10.1016/S0002-9440(10)62976-8) PMID: 16049318
17. Reifenberg K, Lehr HA, Torzewski M, Steige G, Wiese E, Küpper I et al. Interferon-gamma induces chronic active myocarditis and cardiomyopathy in transgenic mice. *Am J Pathol* 2007; 171: 463–472. <https://doi.org/10.2353/ajpath.2007.060906> PMID: 17556594
18. Torzewski M, Wenzel P, Kleinert H, Becker C, El-Masri J, Wiese E et al. Chronic inflammatory cardiomyopathy of interferon gamma-overexpressing transgenic mice is mediated by tumor necrosis factor-alpha. *Am J Pathol* 2012; 180: 73–81. <https://doi.org/10.1016/j.ajpath.2011.09.006> PMID: 22051774
19. Kindermann I, Kindermann M, Kandolf R, Klingel K, Bultmann B, Müller T et al. Predictors of outcome in patients with suspected myocarditis. *Circulation* 2008; 118: 639–648. <https://doi.org/10.1161/CIRCULATIONAHA.108.769489> PMID: 18645053
20. Valaperti A Drugs Targeting the Canonical NF-kappaB Pathway to Treat Viral and Autoimmune Myocarditis. *Curr Pharm Des* 2016; 22: 440–449. <https://doi.org/10.2174/1381612822666151222160409> PMID: 26696258
21. Brown DA, Perry JB, Allen ME, Sabbah HN, Stauffer BL, Shaikh SR et al. Expert consensus document: Mitochondrial function as a therapeutic target in heart failure. *Nat Rev Cardiol* 2017; 14: 238–250. <https://doi.org/10.1038/nrcardio.2016.203> PMID: 28004807
22. Pall ML The NO/ONOO-cycle as the central cause of heart failure. *Int J Mol Sci* 2013; 14: 22274–22330. <https://doi.org/10.3390/ijms141122274> PMID: 24232452
23. Teixeira PC, Santos RH, Fiorelli AI, Bilate AM, Benvenuti LA, Stolf NA et al. Selective decrease of components of the creatine kinase system and ATP synthase complex in chronic Chagas disease cardiomyopathy. *PLoS Negl Trop Dis* 2011; 5: e1205. <https://doi.org/10.1371/journal.pntd.0001205> PMID: 21738806
24. Lopez M, Tanowitz HB, Garg NJ Pathogenesis of Chronic Chagas Disease: Macrophages, Mitochondria, and Oxidative Stress. *Curr Clin Microbiol Rep* 2018; 5: 45–54. PMID: 29868332
25. Yan B, Wang H, Tan Y, Fu W microRNAs in Cardiovascular Disease: Small Molecules but Big Roles. *Curr Top Med Chem* 2019; 19: 1918–1947. <https://doi.org/10.2174/1568026619666190808160241> PMID: 31393249
26. Rupaimoole R, Slack FJ MicroRNA therapeutics: towards a new era for the management of cancer and other diseases. *Nat Rev Drug Discov* 2017; 16: 203–222. <https://doi.org/10.1038/nrd.2016.246> PMID: 28209991
27. Ferreira LRP, Ferreira FM, Laugier L, Cabantous S, Navarro IC, da Silva Candido D et al. Integration of miRNA and gene expression profiles suggest a role for miRNAs in the pathobiological processes of acute *Trypanosoma cruzi* infection. *Sci Rep* 2017; 7: 17990. <https://doi.org/10.1038/s41598-017-18080-9> PMID: 29269773
28. Ferreira LR, Frade AF, Santos RH, Teixeira PC, Baron MA, Navarro IC et al. MicroRNAs miR-1, miR-133a, miR-133b, miR-208a and miR-208b are dysregulated in Chronic Chagas disease Cardiomyopathy. *Int J Cardiol* 2014; 175: 409–417. <https://doi.org/10.1016/j.ijcard.2014.05.019> PMID: 24910366
29. Zhang J, Chen M, Chen J, Lin S, Cai D, Chen C et al. Long non-coding RNA MIAT acts as a biomarker in diabetic retinopathy by absorbing miR-29b and regulating cell apoptosis. *Biosci Rep* 2017; 37(2): BSR20170036. <https://doi.org/10.1042/BSR20170036> PMID: 28246353
30. Zhu XH, Yuan YX, Rao SL, Wang P LncRNA MIAT enhances cardiac hypertrophy partly through sponging miR-150. *Eur Rev Med Pharmacol Sci* 2016; 20: 3653–3660. PMID: 27649667
31. Yang Y, Wu J, Wang J A database and functional annotation of NF-kB target genes. *Int J Clin Exp Med* 2016; 9: 7986–7995.
32. Calvo SE, Clauser KR, Mootha VK MitoCarta2.0: an updated inventory of mammalian mitochondrial proteins. *Nucleic Acids Res* 2016; 44: D1251–1257. <https://doi.org/10.1093/nar/gkv1003> PMID: 26450961
33. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meililes GV et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* 2013; 14: 128. <https://doi.org/10.1186/1471-2105-14-128> PMID: 23586463
34. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 2003; 13: 2498–2504. <https://doi.org/10.1101/gr.1239303> PMID: 14597658

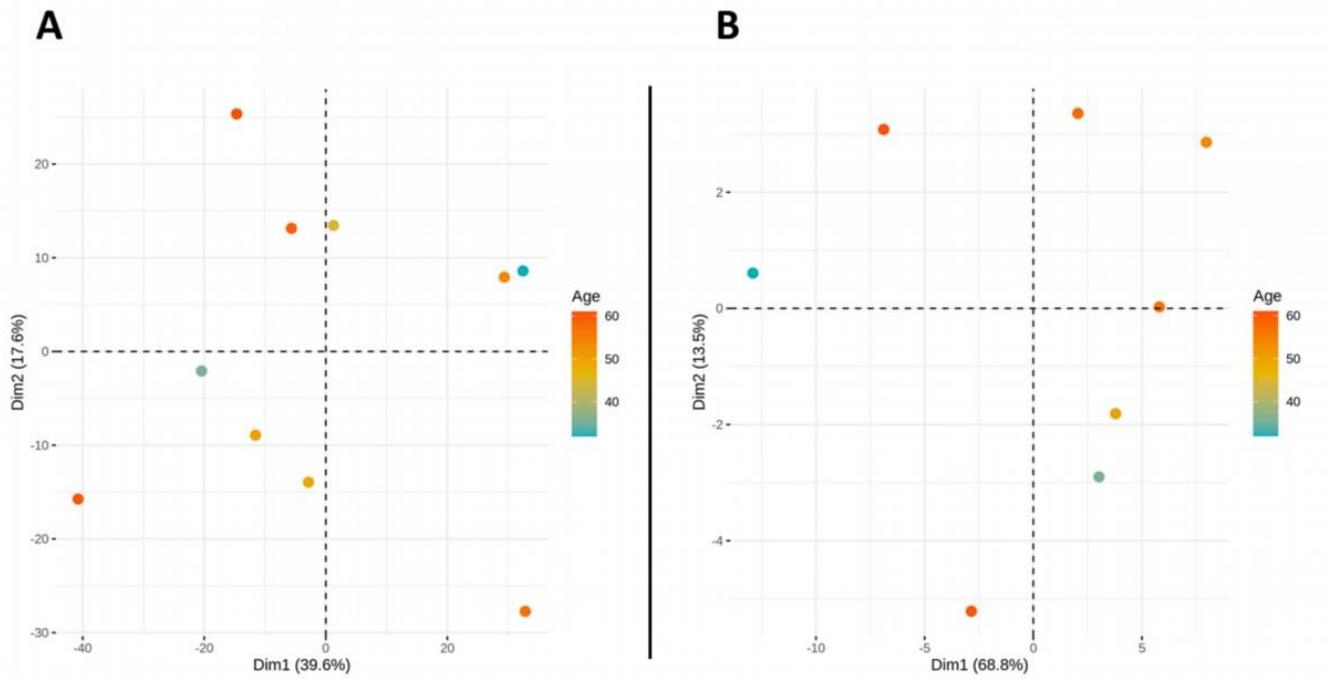
35. Navarro IC, Ferreira FM, Nakaya HI, Baron MA, Vilar-Pereira G, Pereira IR et al. MicroRNA Transcriptome Profiling in Heart of *Trypanosoma cruzi*-Infected Mice: Parasitological and Cardiological Outcomes. *PLoS Negl Trop Dis* 2015; 9: e0003828. <https://doi.org/10.1371/journal.pntd.0003828> PMID: 26086673
36. Dvinge H, Bertone P HTqPCR: high-throughput analysis and visualization of quantitative real-time PCR data in R. *Bioinformatics* 2009; 25: 3325–3326. <https://doi.org/10.1093/bioinformatics/btp578> PMID: 19808880
37. Kauffmann A, Gentleman R, Huber W arrayQualityMetrics—a bioconductor package for quality assessment of microarray data. *Bioinformatics* 2009; 25: 415–416. <https://doi.org/10.1093/bioinformatics/btn647> PMID: 19106121
38. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 2015; 43: e47. <https://doi.org/10.1093/nar/gkv007> PMID: 25605792
39. Laugier L, Frade AF, Ferreira FM, Baron MA, Teixeira PC, Cabantous S et al. Whole-Genome Cardiac DNA Methylation Fingerprint and Gene Expression Analysis Provide New Insights in the Pathogenesis of Chronic Chagas Disease Cardiomyopathy. *Clin Infect Dis* 2017; 65: 1103–1111. <https://doi.org/10.1093/cid/cix506> PMID: 28575239
40. Chae HJ, Ha KC, Kim DS, Cheung GS, Kwak YG, Kim H-M et al. Catalase protects cardiomyocytes via its inhibition of nitric oxide synthesis. *Nitric Oxide* 2006; 14: 189–199. <https://doi.org/10.1016/j.niox.2005.11.008> PMID: 16403660
41. Patten M, Kramer E, Bunemann J, Wenck C, Thoenes M, Long WC Endotoxin and cytokines alter contractile protein expression in cardiac myocytes in vivo. *Pflugers Arch* 2001; 442: 920–927. <https://doi.org/10.1007/s004240100612> PMID: 11680626
42. Wang Z, Jiang B, Brecher P Selective inhibition of STAT3 phosphorylation by sodium salicylate in cardiac fibroblasts. *Biochem Pharmacol* 2002; 63: 1197–1207. [https://doi.org/10.1016/s0006-2952\(02\)00853-5](https://doi.org/10.1016/s0006-2952(02)00853-5) PMID: 11960596
43. Dai B, Cui M, Zhu M, Su WL, Qiu MC, Zhang H STAT1/3 and ERK1/2 synergistically regulate cardiac fibrosis induced by high glucose. *Cell Physiol Biochem* 2013; 32: 960–971. <https://doi.org/10.1159/000354499> PMID: 24107317
44. Wan X, Wen JJ, Koo SJ, Liang LY, Garg NJSIRT1-PGC1 α -NF κ B Pathway of Oxidative and Inflammatory Stress during *Trypanosoma cruzi* Infection: Benefits of SIRT1-Targeted Therapy in Improving Heart Function in Chagas Disease. *PLoS Pathog* 2016; 12: e1005954. <https://doi.org/10.1371/journal.ppat.1005954> PMID: 27764247
45. Wu AJ, Lafrenie RM, Park C, Apinhasmit W, Chen ZJ, Birkedal-Hansen H et al. Modulation of MMP-2 (gelatinase A) and MMP-9 (gelatinase B) by interferon-gamma in a human salivary gland cell line. *J Cell Physiol* 1997; 171: 117–124. [https://doi.org/10.1002/\(SICI\)1097-4652\(199705\)171:2<117::AID-JCP1>3.0.CO;2-R](https://doi.org/10.1002/(SICI)1097-4652(199705)171:2<117::AID-JCP1>3.0.CO;2-R) PMID: 9130458
46. Levick SP, Goldspink PH Could interferon-gamma be a therapeutic target for treating heart failure? *Heart Fail Rev* 2014; 19: 227–236. <https://doi.org/10.1007/s10741-013-9393-8> PMID: 23589353
47. Maier HJ, Schips TG, Wietelmann A, Kruger M, Brunner C, Sauter M et al. Cardiomyocyte-specific I κ B kinase (IKK)/NF- κ B activation induces reversible inflammatory cardiomyopathy and heart failure. *Proc Natl Acad Sci U S A* 2012; 109: 11794–11799. <https://doi.org/10.1073/pnas.1116584109> PMID: 22753500
48. Silva JS, Vespa GN, Cardoso MA, Aliberti JC, Cunha FQ Tumor necrosis factor α mediates resistance to *Trypanosoma cruzi* infection in mice by inducing nitric oxide production in infected gamma interferon-activated macrophages. *Infect Immun* 1995; 63: 4862–4867. <https://doi.org/10.1128/IAI.63.12.4862-4867.1995> PMID: 7591147
49. Vila-del Sol V, Punzon C, Fresno M IFN-gamma-induced TNF- α expression is regulated by interferon regulatory factors 1 and 8 in mouse macrophages. *J Immunol* 2008; 181: 4461–4470. <https://doi.org/10.4049/jimmunol.181.7.4461> PMID: 18802049
50. Cheshire JL, Baldwin AS Jr. Synergistic activation of NF- κ B by tumor necrosis factor α and gamma interferon via enhanced I κ B α degradation and de novo I κ B β degradation. *Mol Cell Biol* 1997; 17: 6746–6754. <https://doi.org/10.1128/mcb.17.11.6746> PMID: 9343439
51. Luss H, Watkins SC, Freeswick PD, Imro AK, Nussler AK, Billiar TR et al. Characterization of inducible nitric oxide synthase expression in endotoxemic rat cardiac myocytes in vivo and following cytokine exposure in vitro. *J Mol Cell Cardiol* 1995; 27: 2015–2029. [https://doi.org/10.1016/0022-2828\(95\)90023-3](https://doi.org/10.1016/0022-2828(95)90023-3) PMID: 8523461
52. Wang D, McMillin JB, Bick R, Buja Response of the neonatal rat cardiomyocyte in culture to energy depletion: effects of cytokines, nitric oxide, and heat shock proteins. *Lab Invest* 1996; 75: 809–818. PMID: 8973476

53. Reinsbach S, Nazarov PV, Philippidou D, Schmitt M, Wienecke-Baldacchino A, Muller A et al. Dynamic regulation of microRNA expression following interferon-gamma-induced gene transcription. *RNA Biol* 2012; 9: 978–989. <https://doi.org/10.4161/rna.20494> PMID: 22767256
54. Zhao AQ, Xie H, Lin SY, Lei Q, Ren WX, Gao F et al. Interferon-gamma alters the immune-related miRNA expression of microvesicles derived from mesenchymal stem cells. *J Huazhong Univ Sci Technol Med Sci* 2017; 37: 179–184. <https://doi.org/10.1007/s11596-017-1712-1> PMID: 28397044
55. Tijssen AJ, van der Made I, van den Hoogenhof MM, Wijnen WJ, van Deel ED, De Groot NE et al. The microRNA-15 family inhibits the TGFbeta-pathway in the heart. *Cardiovasc Res* 2014; 104: 61–71. <https://doi.org/10.1093/cvr/cvu184> PMID: 25103110
56. Rawal S, Munasinghe PE, Nagesh PT, Lew JKS, Jones GT, Williams MJA et al. Down-regulation of miR-15a/b accelerates fibrotic remodelling in the Type 2 diabetic human and mouse heart. *Clin Sci (Lond)* 2017; 131: 847–863. <https://doi.org/10.1042/CS20160916> PMID: 28289072
57. Liu Y, Taylor NE, Lu L, Usa K, Cowley AW Jr., Ferreri NR et al. Renal medullary microRNAs in Dahl salt-sensitive rats: miR-29b regulates several collagens and related genes. *Hypertension* 2010; 55: 974–982. <https://doi.org/10.1161/HYPERTENSIONAHA.109.144428> PMID: 20194304
58. Sengupta S, den Boon JA, Chen IH, Newton MA, Stanhope SA, Cheng Y-J et al. MicroRNA 29c is down-regulated in nasopharyngeal carcinomas, up-regulating mRNAs encoding extracellular matrix proteins. *Proc Natl Acad Sci U S A* 2008; 105: 5874–5878. <https://doi.org/10.1073/pnas.0801130105> PMID: 18390668
59. van Rooij E, Sutherland LB, Liu N, Williams AH, McAnally J, Gerard RD et al. A signature pattern of stress-responsive microRNAs that can evoke cardiac hypertrophy and heart failure. *Proc Natl Acad Sci U S A* 2006; 103: 18255–18260. <https://doi.org/10.1073/pnas.0608791103> PMID: 17108080
60. Frade AF, Laugier L, Ferreira LR, Baron MA, Benvenuti LA, Teixeira PC et al. Myocardial Infarction-Associated Transcript, a Long Noncoding RNA, Is Overexpressed During Dilated Cardiomyopathy Due to Chronic Chagas Disease. *J Infect Dis* 2016; 214: 161–165. <https://doi.org/10.1093/infdis/jiw095> PMID: 26951817
61. Ellis KL, Cameron VA, Troughton RW, Frampton CM, Ellmers LJ, Richard AM Circulating microRNAs as candidate markers to distinguish heart failure in breathless patients. *Eur J Heart Fail* 2013; 15: 1138–1147. <https://doi.org/10.1093/eurjhf/hft078> PMID: 23696613
62. Qi H, Ren J, E M, Zhang Q, Cao Y, Ba L et al. MiR-103 inhibiting cardiac hypertrophy through inactivation of myocardial cell autophagy via targeting TRPV3 channel in rat hearts. *J Cell Mol Med* 2019; 23: 1926–1939. <https://doi.org/10.1111/jcmm.14095> PMID: 30604587
63. Wang X, Ha T, Zou J, Ren D, Liu L, Zhang X et al. MicroRNA-125b protects against myocardial ischaemia/reperfusion injury via targeting p53-mediated apoptotic signalling and TRAF6. *Cardiovasc Res* 2014; 102: 385–395. <https://doi.org/10.1093/cvr/cvu044> PMID: 24576954
64. Li S, Zhu J, Zhang W, Chen Y, Zhang K, Popescu LM et al. Signature microRNA expression profile of essential hypertension and its novel link to human cytomegalovirus infection. *Circulation* 2011; 124: 175–184. <https://doi.org/10.1161/CIRCULATIONAHA.110.012237> PMID: 21690488
65. Cheng X, Ku CH, Siow RC Regulation of the Nrf2 antioxidant pathway by microRNAs: New players in micromanaging redox homeostasis. *Free Radic Biol Med* 2013; 64: 4–11. <https://doi.org/10.1016/j.freeradbiomed.2013.07.025> PMID: 23880293
66. Soares MP, Gozzelino R, Weis STissue damage control in disease tolerance. *Trends Immunol* 2014; 35: 483–494. <https://doi.org/10.1016/j.it.2014.08.001> PMID: 25182198
67. Soares MP, Ribeiro AM Nrf2 as a master regulator of tissue damage control and disease tolerance to infection. *Biochem Soc Trans* 2015; 43: 663–668. <https://doi.org/10.1042/BST20150054> PMID: 26551709
68. Demkes CJ, van Rooij E MicroRNA-146a as a Regulator of Cardiac Energy Metabolism. *Circulation* 2017; 136: 762–764. <https://doi.org/10.1161/CIRCULATIONAHA.117.029703> PMID: 28827220
69. Bersell K, Arab S, Haring B, Kuhn B Neuregulin1/ErbB4 signaling induces cardiomyocyte proliferation and repair of heart injury. *Cell* 2009 138: 257–270. <https://doi.org/10.1016/j.cell.2009.04.060> PMID: 19632177
70. Horie T, Ono K, Nishi H, Nagao K, Kinoshita M, Watanabe S et al. Acute doxorubicin cardiotoxicity is associated with miR-146a-induced inhibition of the neuregulin-ErbB pathway. *Cardiovasc Res* 2010; 87: 656–664. <https://doi.org/10.1093/cvr/cvq148> PMID: 20495188
71. Gao M, Wang X, Zhang X, Ha T, Ma H, Liu L et al. Attenuation of Cardiac Dysfunction in Polymicrobial Sepsis by MicroRNA-146a Is Mediated via Targeting of IRAK1 and TRAF6 Expression. *J Immunol* 2015; 195: 672–682. <https://doi.org/10.4049/jimmunol.1403155> PMID: 26048146

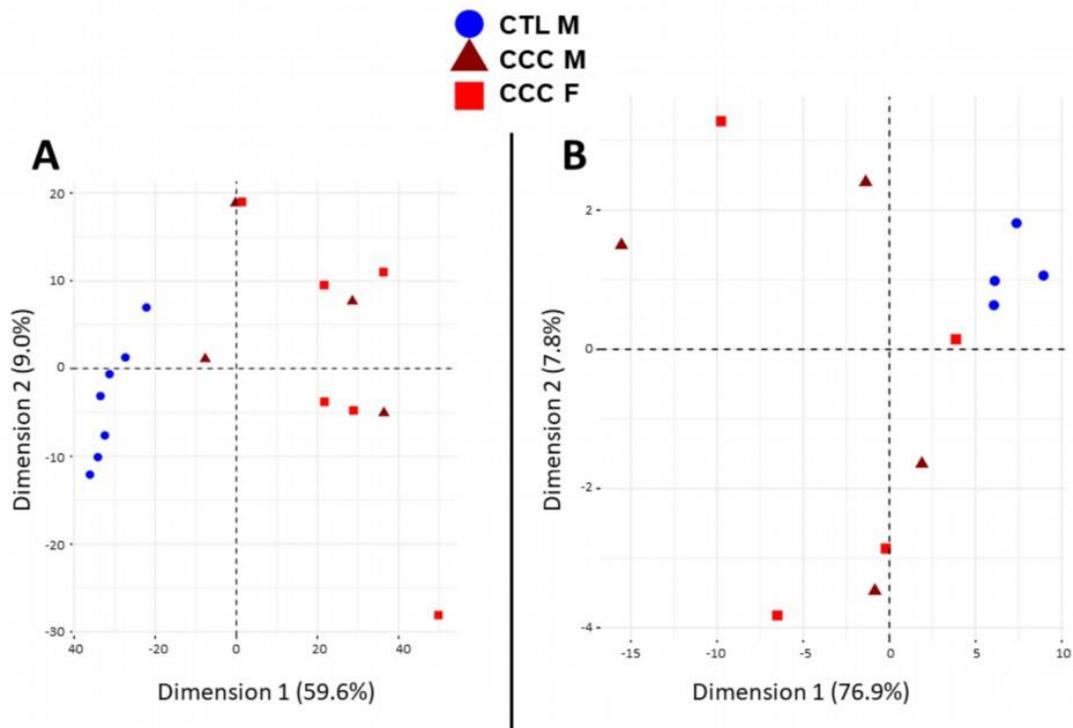
72. Taganov KD, Boldin MP, Chang KJ, Baltimore D NF-kappaB-dependent induction of microRNA miR-146, an inhibitor targeted to signaling proteins of innate immune responses. *Proc Natl Acad Sci U S A* 2006; 103: 12481–12486. <https://doi.org/10.1073/pnas.0605298103> PMID: 16885212
73. Lu LF, Boldin MP, Chaudhry A, Lin LL, Taganov KD, Hanada T et al. Function of miR-146a in controlling Treg cell-mediated regulation of Th1 responses. *Cell* 2010; 142: 914–929. <https://doi.org/10.1016/j.cell.2010.08.012> PMID: 20850013
74. Ramos GC, van den Berg A, Nunes-Silva V, Weirather J, Peters L, Burkard M et al. Myocardial aging as a T-cell-mediated phenomenon. *Proc Natl Acad Sci U S A* 2017; 114: E2420–E2429. <https://doi.org/10.1073/pnas.1621047114> PMID: 28255084

A**B**

Supplementary figure 1. Unsupervised hierarchical clustering done on patients with severe chronic Chagas disease cardiomyopathy (CCC) and controls. **A.** Unsupervised hierarchical clustering based on the 1535 DEGs. **B.** Unsupervised hierarchical clustering based on the 80 DEMs.



Supplementary figure 2. Principal component analysis (PCA) plots taking into account the age of the cases. Principal component analysis (PCA) plot of samples was performed based **A.** on 1535 differentially expressed genes (DEGs) between CCC and controls. **B.** on 80 differentially expressed miRNAs (DEMs) between CCC and controls. Each plot was generated only on cases and the age of the patients was overlaid.



Supplementary figure 3. Principal component analysis (PCA) plots taking into account the sex of the cases and controls. Principal component analysis (PCA) plot of samples was performed based **A.** on 1535 differentially expressed genes (DEGs) between CCC and controls. **B.** on 80 differentially expressed miRNAs (DEMs) between CCC and controls. On each plot cases and controls are indicated according to their sex.

Toutes les tables supplémentaires sont directement **accessibles en ligne** à l'adresse suivante :

<https://journals.plos.org/plosntds/article?id=10.1371/journal.pntd.0008889>

2.2. La méthylation de l'ADN est impliquée dans le développement des cardiomyopathies chagasiennes chroniques, et semble réguler l'expression de gènes en affectant la fixation de facteurs de transcription

2.2.1. Implication de la méthylation sur la régulation génique et les processus biologiques dans les cardiomyopathies chagasiennes chroniques

Article 3 : Epigenetic regulation of transcription factor binding motifs promotes Th1 response in Chagas disease cardiomyopathy.

Pauline Brochet, Barbara Maria Ianni, Laurie Laugier, Amanda Farage Frade, João Paulo Silva Nunes, Priscila Camillo Teixeira, Charles Mady, Ludmila Rodrigues Pinto Ferreira, Quentin Ferré, Ronaldo Honorato Barros Santos, Andreia Kuramoto, Sandrine Cabantous, Samuel Steffen, Antonio Noedir Stolf, Pablo Pomerantzeff, Alfredo Inacio Fiorelli, Edimar Alcides Bocchi, Cristina Wide Pissetti, Bruno Saba, Darlan da Silva Cândido, Fabrício C Dias, Marcelo Ferraz Sampaio, Fabio Antônio Gaiotto, José Antonio Marin-Neto, Abílio Fragata, Ricardo Costa Fernandes Zaniratto, Sergio Siqueira, Giselle De Lima Peixoto, Vagner Oliveira-Carvalho Rigaud, Fernando Bacal, Paula Buck, Rafael Ribeiro Almeida, Hui Tzu Lin-Wang, André Schmidt, Martino Martinelli, Mario Hiroyuki Hirata, Eduardo Antonio Donadi, Alexandre Costa Pereira, Virmondes Rodrigues Junior, Denis Puthier, Jorge Kalil, Lionel Spinelli*, Edecio Cunha-Neto*, Christophe Chevillard*.

Front Immunol. 2022 Aug 22;13:958200. doi: 10.3389/fimmu.2022.958200. PMID: 36072583; PMCID: PMC9441916.

Présentation de la publication :

L'impact de la méthylation dans les CCC avait déjà été abordé dans une précédente étude, mais était limité par des contraintes techniques (450.000 CpG analysés), de même que les variations génétiques mentionnées précédemment (analysées avec une puce à ADN, 27.958 gènes cibles mais seulement 7.419 ARNInc).

Afin d'approfondir ces études, nous avons à nouveau analysé le transcriptome et le méthylome à partir de tissu cardiaque de 6 individus sains et 8 patients atteints de CCC, avec les méthodes de RNA-sequencing et Illumina EPIC methylation (850.000 CpG couvrant 96% des gènes). De plus, l'expression génique dérivant du tissu cardiaque de 8 patients atteints de cardiomyopathies dilatées (CMD) a été intégrée à l'analyse. Les données ont été alignées contre le génome de référence hg19 avec STAR, et l'analyse des gènes a été menée avec DESeq2. Les gènes présentant une valeur absolue du $\log_2(\text{FC}) \geq 1,5$ et une p-valeur corrigée $\leq 0,05$ ont été identifiés comme différentiellement exprimés. Un total de 1407 gènes a été significativement associés aux CCC (3.23% du nombre total de gènes), majoritairement sur-exprimés dans la pathologie. Seuls 9% de ces gènes DE ont été retrouvés en commun avec ceux propres aux CMD. Alors que les CMD présentaient des altérations de l'expression génique associées à la régulation des cellules musculaires lisses ou aux processus apoptotiques, les CCC manifestaient clairement un profil inflammatoire, avec une réponse associée aux lymphocytes T, à l'interféron gamma, à la neuroinflammation ainsi qu'à la régulation de la production et de la présentation d'antigènes. En revanche, les deux cardiomyopathies exprimaient des perturbations génétiques dans le transport des ions, notamment calcium. Ces résultats sont cohérents avec l'infiltrat immunitaire présent dans le cœur des CCC, évalué lors de cette étude par une analyse de déconvolution. En effet, les CCC ont affiché une plus grande proportion de lymphocytes CD8⁺, mémoires et folliculaires, ainsi que de macrophages M2 et de cellules NK que les contrôles sains et les CMD. Parallèlement, la proportion de cardiomyocytes et de cellules musculaires lisses a été observée significativement réduite.

Une étude détaillée des ARN non codants a permis d'associer, grâce à plusieurs bases de données publiques, 19 miARN et 145 ARNInc aux CCC, dont 143 non retrouvés chez les CMD. 6 de ces ARNInc sont connus pour cibler 4 gènes et 2 miARN différents, comme MIAT et miR-133a, ou MIR155HG et miR155. Ces cibles, de même que MIAT, ont par ailleurs déjà été fortement associées aux CCC.

Concernant la méthylation, 16,883 CpG différentiellement méthylés (valeur absolue de bêta $\geq 0,2$ et p-valeur corrigée $\leq 0,05$) ont été identifiés grâce à la librairie R ChAMP, soit 2,34% de l'ensemble des CpG. Ces derniers correspondent à 5,814 gènes. En particulier, 996 CpG sont localisés dans 390 gènes DE. Les régions promotrices de ces gènes (définies ici comme incluant les 1500 bases avant le site d'initiation de la transcription, la région 5'UTR et le 1er exon) sont par ailleurs significativement enrichies en sites de méthylation différentiellement méthylés.

Nous avons affiné notre analyse et détecté 92 régions différentiellement méthylées entre les CCC et les contrôles localisées dans les régions régulatrices de 89 gènes DE. En utilisant l'outil Ologram et la base de données ReMAP, nous avons prédit 30 facteurs de transcription (FT) se fixant préférentiellement sur ces régions. En se basant sur la base de données Jaspar et en utilisant l'outil Fimo, nous avons confirmé que les motifs de fixation de ces FT correspondent significativement aux régions d'ADN différentiellement méthylés. Les 30 FT identifiés sont impliqués dans la réponse immunitaire, la régulation de la croissance du muscle cardiaque ou le développement des neurones du système nerveux périphérique. L'étude détaillée des gènes associés à la réponse lymphocyte Th1 a mis en avant 28 FT, appuyant leur impact dans la réponse immunitaire exacerbée chez les CCC. Parmi eux, RUNX3 et TBX21, déjà trouvés chez les CCC, semblent avoir une grande importance, de même que certains nouveaux candidats (BRD4, BCLAF1, CBF1 et PAX5). La pertinence de TBX21 et RUNX3 dans la pathogénèse des CCC a été particulièrement relevée, leur motif de fixation étant enrichi dans la majorité des gènes DE. RUNX3, ayant une région différentiellement méthylée dans son promoteur, pourrait également être la cible de différents facteurs de transcription, incluant lui-même ainsi que PAX5, YY1, SP1, MAX, EBF1 et IRF4. De plus, une analyse combinant l'implication combinée de plusieurs FT a montré que PAX5 et IRF4 pourraient interagir ensemble. Ces résultats illustrent l'importance de la méthylation dans la régulation de l'expression génétique dans les CCC, mais également la complexité de ces mécanismes.

Pour finir, une étude de la méthylation a également été réalisée sur le sang de patients atteints de formes chroniques, notamment asymptomatiques (n = 48), CCC modérée (n = 48) ou CCC sévère (n = 48). 12.624 CpG varient entre les formes asymptomatiques et symptomatiques, et 6.735 sont impliqués dans le développement des CCC. Bien qu'un faible nombre de gènes ait été retrouvé en commun entre toutes ces analyses (transcriptomique, méthylomique sur le tissu et méthylomique sur le sang), les processus biologiques affectés sont similaires. Toutefois, en se concentrant sur l'évolution des CCC, la majorité des gènes sont impliqués dans la neurogénèse, le système cardiovasculaire ou l'organisation des filaments d'actine, illustrant bien l'aggravation de la maladie, sans changement majeur de la réponse immune.

Globalement, cette étude a permis de confirmer l'implication de la méthylation dans des processus déjà connus chez les CCC, et d'affiner nos connaissances sur l'évolution de la pathologie.

De plus, elle a permis de mettre en évidence le rôle de la méthylation dans la régulation génétique, et d'identifier des facteurs de transcriptions qui pourraient contribuer à la pathogenèse de cette maladie.

Contribution personnelle :

Ma contribution dans cet article, dont je suis la première auteur, a été majoritaire. J'ai pris en charge l'intégralité des analyses bioinformatiques, depuis le traitement des données brutes jusqu'aux études statistiques. J'ai également rédigé intégralement la première version de l'article, et ai participé aux corrections.



OPEN ACCESS

EDITED BY

Steven O'Reilly,
STipe Therapeutics, Denmark

REVIEWED BY

Miguel Vicco,
GlaxoSmithKline, Belgium
Francisco Callejas Hernández,
New York University, United States
Jose Roberto Sotelo-Silveira,
Instituto de Investigaciones Biológicas
Clemente Estable (IIBCE),
Uruguay

*CORRESPONDENCE

Christophe Chevillard
christophe.chevillard@univ-amu.fr
Edecio Cunha-Neto
edecunha@gmail.com
Lionel Spinelli
lionel.spinelli@univ-amu.fr†These authors have contributed
equally to this work

SPECIALTY SECTION

This article was submitted to
Inflammation,
a section of the journal
Frontiers in Immunology

RECEIVED 31 May 2022

ACCEPTED 26 July 2022

PUBLISHED 22 August 2022

CITATION

Brochet P, Ianni BM, Laugier L,
Frade AF, Silva Nunes JP, Teixeira PC,
Mady C, Ferreira LRP, Ferré Q,
Santos RHB, Kuramoto A,
Cabantous S, Steffen S, Stolf AN,
Pomerantzeff P, Fiorelli AI, Bocchi EA,
Pissetti CW, Saba B, Cândido DdS,
Dias FC, Sampaio MF, Gaiotto FA,
Marin-Neto JA, Fragata A,
Zaniratto RCF, Siqueira S, Peixoto GDL,
Rigaud VO-C, Bacal F, Buck P,
Almeida RR, Lin-Wang HT, Schmidt A,
Martinelli M, Hirata MH, Donadi EA,
Costa Pereira A, Rodrigues Junior V,
Puthier D, Kalil J, Spinelli L,
Cunha-Neto E and Chevillard C (2022)
Epigenetic regulation of transcription
factor binding motifs promotes Th1
response in Chagas
disease cardiomyopathy.
Front. Immunol. 13:958200.
doi: 10.3389/fimmu.2022.958200

Epigenetic regulation of transcription factor binding motifs promotes Th1 response in Chagas disease cardiomyopathy

Pauline Brochet¹, Barbara Maria Ianni², Laurie Laugier³,
Amanda Farage Frade^{2,4,5}, João Paulo Silva Nunes^{1,2,4,5},
Priscila Camillo Teixeira^{2,4,5}, Charles Mady⁶,
Ludmila Rodrigues Pinto Ferreira⁷, Quentin Ferré¹,
Ronaldo Honorato Barros Santos⁸, Andreia Kuramoto²,
Sandrine Cabantous³, Samuel Steffen^{8,9}, Antonio Noedir Stolf⁹,
Pablo Pomerantzeff¹⁰, Alfredo Inacio Fiorelli⁹,
Edimar Alcides Bocchi⁹, Cristina Wide Pissetti¹¹, Bruno Saba¹²,
Darlan da Silva Cândido^{2,4,5}, Fabrício C. Dias¹³,
Marcelo Ferraz Sampaio¹², Fabio Antônio Gaiotto^{8,9},
José Antonio Marin-Neto¹³, Abílio Fragata¹²,
Ricardo Costa Fernandes Zaniratto², Sergio Siqueira¹⁴,
Giselle De Lima Peixoto¹⁴, Vagner Oliveira-Carvalho Rigaud^{2,15},
Fernando Bacal⁸, Paula Buck¹⁰, Rafael Ribeiro Almeida^{2,4,5},
Hui Tzu Lin-Wang¹², André Schmidt¹³, Martino Martinelli¹⁴,
Mario Hiroyuki Hirata¹⁶, Eduardo Antonio Donadi¹³,
Alexandre Costa Pereira¹⁰, Virmondos Rodrigues Junior¹¹,
Denis Puthier¹, Jorge Kalil^{2,4,5}, Lionel Spinelli^{1,17*},
Edecio Cunha-Neto^{2,4,5*†} and Christophe Chevillard^{1*†}

¹Institut National de la Santé Et de la Recherche Médicale (INSERM), Unité Mixte de Recherche (UMR) _1090, Aix Marseille Université, TAGC Theories and Approaches of Genomic Complexity, Institut MarMaRa, Marseille, France, ²Laboratory of Immunology, Heart Institute Instituto do Coração (InCor), University of São Paulo, School of Medicine, São Paulo, Brazil, ³Aix Marseille Université, Génétique et Immunologie des Maladies Parasitaires, Inserm, UMR_906, Marseille, France, ⁴Division of Clinical Immunology and Allergy, University of São Paulo, School of Medicine, São Paulo, Brazil, ⁵Instituto Nacional de Ciência e Tecnologia, INCT, III- Institute for Investigation in Immunology, São Paulo, Brazil, ⁶Myocardopathies and Aortic Diseases Unit, Heart Institute Instituto do Coração (InCor), School of Medicine, University of São Paulo, São Paulo, Brazil, ⁷RNA Systems Biology Laboratory (RSBL), Departamento de Morfologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil, ⁸Division of Transplantation, Heart Institute Instituto do Coração (InCor), University of São Paulo, School of Medicine, São Paulo, Brazil, ⁹Division of Surgery, Heart Institute Instituto do Coração (InCor), University of São Paulo, School of Medicine, São Paulo, Brazil, ¹⁰Heart Institute (InCor), School of Medicine, University of São Paulo, São Paulo, São Paulo, Brazil, ¹¹Laboratory of Immunology, Universidade Federal Do Triângulo Mineiro (UFMT), Uberaba, Brazil, ¹²Laboratório de Investigação Molecular em Cardiologia, Instituto de Cardiologia Dante Pazzanese (IDPC), São Paulo, Brazil, ¹³School of Medicine of Ribeirão Preto Faculdade de Medicina de Ribeirão Preto (FMRP), University of São Paulo, Ribeirão Preto, Brazil, ¹⁴Pacemaker Clinic, Heart Institute Instituto do Coração (InCor), School of Medicine, University of São Paulo, São Paulo, Brazil, ¹⁵Heart Failure Unit, Heart Institute Instituto do Coração (InCor) School of Medicine, University of Sao Paulo, Sao Paulo, Brazil, ¹⁶Department of Clinical and Toxicological Analyses, Faculty of Pharmaceutical Sciences, University of São Paulo (USP), São Paulo, Brazil, ¹⁷Aix Marseille Université, CNRS, INSERM, Centre d'Immunologie de Marseille-Luminy, Marseille, France

Chagas disease, caused by the protozoan *Trypanosoma cruzi*, is an endemic parasitic disease of Latin America, affecting 7 million people. Although most patients are asymptomatic, 30% develop complications, including the often-fatal Chronic Chagasic Cardiomyopathy (CCC). Although previous studies have demonstrated some genetic deregulations associated with CCCs, the causes of their deregulations remain poorly described. Based on bulk RNA-seq and whole genome DNA methylation data, we investigated the genetic and epigenetic deregulations present in the moderate and severe stages of CCC. Analysis of heart tissue gene expression profile allowed us to identify 1407 differentially expressed transcripts (DEGs) specific from CCC patients. A tissue DNA methylation analysis done on the same tissue has permitted the identification of 92 regulatory Differentially Methylated Regions (DMR) localized in the promoter of DEGs. An in-depth study of the transcription factors binding sites (TFBS) in the DMRs corroborated the importance of TFBS's DNA methylation for gene expression in CCC myocardium. TBX21, RUNX3 and EBF1 are the transcription factors whose binding motif appears to be affected by DNA methylation in the largest number of genes. By combining both transcriptomic and methylomic analysis on heart tissue, and methylomic analysis on blood, 4 biological processes affected by severe CCC have been identified, including immune response, ion transport, cardiac muscle processes and nervous system. An additional study on blood methylation of moderate CCC samples put forward the importance of ion transport and nervous system in the development of the disease.

KEYWORDS

dilated cardiomyopathy, Chagas disease, epigenetic, methylation, Th1 response, transcription factors

Introduction

Chagas disease is a neglected disease caused by the protozoan *Trypanosoma cruzi*. This parasite is endemic in 21 Latin America countries, where it affects around 7 million people through an insect vector, Reduviidae. With migratory flows, this disease can now be found in non-endemic countries (1). The clinical course of the disease comprises an acute phase, mostly asymptomatic, and a chronic phase, where 60% of the patients remain asymptomatic. However, 40% develop symptomatic disease, being 10% megaesophagus/megacolon, and 30% Chagas disease cardiomyopathy (CCC) with varying degrees of severity including refractory heart failure (1). This cardiomyopathy is the main cause of deaths from Chagas disease itself and is one of the most lethal cardiomyopathies (2). Some drugs are effective on *T. cruzi* only during the acute phase, but several side effects have been reported (3). The fact that the biological processes leading to CCC are not yet well understood has impaired the development of efficient therapeutic strategies.

The CCC myocardium displays a diffuse myocarditis with signs of inflammatory infiltrate and heart fiber damage, including significant fibrosis. The inflammatory infiltrate of CCC heart lesions is mainly composed of T cells displaying a Th1-like cytokine profile (4). This exacerbated Th1 response is characterized by a high expression of interferon-gamma (IFN- γ), tumor necrosis alpha (TNF- α) and TBX21 (T-bet) (5). Moreover, our group has previously demonstrated that CCC myocardium presents a unique gene expression profile, distinct from the other dilated cardiomyopathies (5, 6).

Many studies have highlighted the importance of DNA methylation in the regulation of gene expression in dilated cardiomyopathy (7), in particular by the methylation/demethylation of transcription factor binding site (TFBS) located in genes regulatory regions (8). Development of severe CCC is also dependent of epigenetic regulations such as DNA methylation (9), but also involving miRNAs (6, 10) and lncRNAs (11). To get a more complete picture of the epigenomic landscape of CCC myocardium, we performed gene expression analysis (RNA-seq) complemented with a

methylation analysis (MethylationEPIC), covering 96% of gene loci, including lncRNA.

Methods

Ethical considerations

The protocol was approved by the institutional review boards of the University of São Paulo School of Medicine and INSERM (French National Institute of Health and Medical Research). Written informed consent was obtained from all patients. All experimental methods comply with the Helsinki Declaration.

Patients and myocardial tissue collection

Human left ventricular free wall heart tissue samples were obtained from patients with end-stage heart failure CCC at the time of heart transplantation (n=8). CCC patients underwent a serological diagnosis of *T. cruzi* infection and standard electrocardiography and echocardiography (12). Biopsies from controls (n=6) were obtained from healthy hearts of organ donors having no suitable recipient, and biopsies for dilated cardiomyopathy (DCM) from end-stage patients (n=8) (Supplementary Table 1).

RNA extraction and sequencing

Total RNAs were extracted from heart tissue samples as previously described (9). Ribosomal RNAs were depleted, and samples were prepared for sequencing according to the Illumina TruSeq RNA Preparation Kit and subjected to pairwise sequencing (2x150bp) with an Illumina HiSeq sequencer. This strategy allows to have information on only the most abundant non-coding RNAs. The RNA-seq data are available under the reference : (GEO accession: GSE191081).

Quality control and alignment

Raw data quality was verified with FastQC (v0.11.5) and reads were filtered removing the adaptors and low-quality based using Trimmomatic (v0.39). Reads were aligned in paired-end mode on GRCh37 (hg19) human reference genome using STAR (v2.5.4b), and gene quantification was done with featureCounts (v2.0.0). All bioinformatic analyses are available at <https://github.com/pbrochet/epiChagas>.

Differential expression analysis

Statistical analyses were performed using R (3.6.2). The *DESeq2* package (v1.26.0) was used for data normalization and differential gene expression analyses (13), using shrinkage function to correct log₂ fold change (log₂(FC)). Benjamini-Hochberg method was applied to obtain False Discovery Rate (FDR) for each analysis. Genes with an FDR ≤ 0.05 and an absolute log₂(FC) greater than 1.5 were considered as differentially expressed (DEG).

Functional enrichment

Gene Ontology Biological Process annotations enrichment was performed using ClueGO cytoscape plugin and KEGG pathway analysis was done with and GAGE (v2.36.0) (14) and pathview R package (v1.26.0). ncRNAs enrichment was realized with three databases: LncRNA2Target, LncTarD and LncRNADisease.

Evaluation of cell types in heart tissue

RNAseq deconvolution was performed using ADAPTS R package. Using data coming from heart tissue microarray and PBMC single-cell RNA-seq datasets, signature matrices were generated with ADAPTS. A Wilcoxon test (FDR<0.05) was applied between healthy control and CCC.

Tissue DNA methylation analysis

DNAs, extracted from whole tissue, were bisulfite converted and amplified with elongation of primers. Amplified DNAs were fragmented and hybridized in EPIC beads according to the protocol described by the manufacturer (Illumina, San Diego, California). Analysis of DNA methylation data was performed with the ChAMP package, using BMIQ for normalization and ComBat for batch-effect correction. Only genomic positions with an FDR<0.05 and a |Δβ|>0.2 were selected. A DMP (Differentially Methylated Position) is associated with a gene when the DMP is inside the gene body or in its promoter region (from gene TSS to 1.5kb upstream). The methylation data are available under the reference : (GEO accession: GSE191082).

Transcription factor binding site characterization

An analysis of differentially methylated regions (DMR) was done with the ChAMP package, using the DMRCate method, with parameters lambda=400 and C=2. A DMR of interest was

defined as a region containing at least 1 DMP located in TSS ([TSS - 1500bp; TSS]), 1st Exon or 5'UTR region of DEGs; and having an FDR ≤ 0.05 . In order to identify transcription factor binding sites (TFBS) affected by a difference in methylation, the ReMap database was used (15). A total of 84 cell lines were selected (Supplementary Table 2), containing immune and heart-related cells, and including 151 transcription factors (TF). First, the transcription factors specifically associated with DMRs were identified with the OLOGRAM tool (16, 17) in a pairwise analysis, meaning we identify the individual TFs enriched with the DMRs. Only those with an FDR ≤ 0.05 were retained. In a second step, we studied the combinations of those selected TFs that were observed in the DMR by using the n-wise overlap option of OLOGRAM (option -more-bed-multiple-overlap, $n \leq 4$). Finally, the known TFBS profiles of identified transcription factor were retrieved from JASPAR database. The location of each TFBS in DMR sequence was identified using FIMO tool.

Blood DNA collection and DNA methylation analysis

Blood (5 to 15 ml of blood) from 96 CCC patients (48 moderate CCC (left ventricular ejection fraction >40%) and 48 severe CCC (left ventricular ejection fraction <40%)) and 48 asymptomatic Chagas disease controls was also collected in EDTA tubes (Supplementary Table 1). Genomic DNA was isolated using standard salted methods and the methylation analysis was done using the same protocol as tissue DNAs.

Results

Headings disregulated genes associated to severe CCC

Gene expression analysis was conducted on left ventricular free wall myocardial tissue from 8 severe CCC patients and 6 healthy organ donors (see workflow: Figure 1). For each sample, we got between 40 and 75 million sequencing reads. Sequences were aligned to the human reference genome GRCh37/hg19. The average mappable rate of the raw reads reached 90% (+/- 2%). No parasite RNA was detected. This is not surprising, since in chronic and end stage patients, the parasite ARN is no longer detected. Gene expression data were obtained from 43533 transcripts. A small fraction of these transcripts (1407/43533 (3.23%)) were considered as differentially expressed between control and CCC (Supplementary Table 3), most part being up-regulated (Supplementary Figure 1A). A specific enrichment occurs in protein coding and non-coding genes, as miRNAs (Supplementary Table 4). PCA and HCA (Hierarchical Clustering Analysis) analysis confirmed that CCC myocardial

gene expression patterns were substantially different from controls (Figure 2A and Supplementary Figure 2A). The sex and the age of the patients have no impact on this clustering (Supplementary Figures 2B, C).

Severe CCC is characterized by a strong inflammatory signature not present in DCM

To understand the pathogenic processes driving CCC, a comparative study was conducted between CCC patients and DCM patients. Among the 3188 genes are differentially expressed in DCM, mostly down-regulated (Supplementary Figure 1B), only 290 (9%) DEGs are in common with CCCs. (Supplementary Table 5). This further emphasizes that the mechanisms involved in CCC and DCM are not similar.

A functional analysis of the DEGs identified both in CCCs or DCMs was conducted (Figure 3 and Supplementary Table 6). Some biological processes are shared by both diseases, like smooth muscle, ERK1/ERK2 cascade and ion transport. Interestingly, calcium ions are particularly affected in CCC, not in DCM. Regarding CCCs, DEGs are almost exclusively specific to the immune response (innate or adaptive). Most of enriched terms are related to T lymphocytes. More specifically, the T CD8+ and T CD4+ are associated to CCC, as well as Th1 response. Besides T cells, other immune cells seem to act in the pathogenic process of CCC, such as B cells, macrophages or NK cells. The regulation of interleukin production is also affected, included IL-1, IL-4, IL-6, IL-10 and IL-12. The chemokine CXCL9 was the most highly expressed among the cytokines (FC=78), followed by CCL19, IFN-gamma, CCL4, CXCL10, CCL17, CCL22, CXCL11, IL-26, LTA, IL-16 and IL-18. Together, these results provide an overview of the pathogenic process associated with CCCs, which is mostly related to the immune response. Furthermore, calcium ion transport seems to be particularly important in CCCs, compared to DCMs.

Several non-coding RNA are specific to severe CCC

Non-coding RNAs are among the elements that can lead to genetic dysregulation. As gene expression analysis was performed by sequencing of total RNAs, we got information on non-coding RNAs (lincRNA, miRNA, snRNA, snoRNA, 3'overlapping ncRNA). Among the 3777 non-coding elements detected in our samples, 179, including 19 miRNAs and 145 lincRNAs; were differentially expressed. Those ncRNAs (Supplementary Figure 3A) were enough to classify samples according to their phenotype, demonstrating the importance of these non-coding elements in chronic Chagas cardiomyopathy. Only 6 had a known target (lincRNA-targeted gene: MIAT-miR-

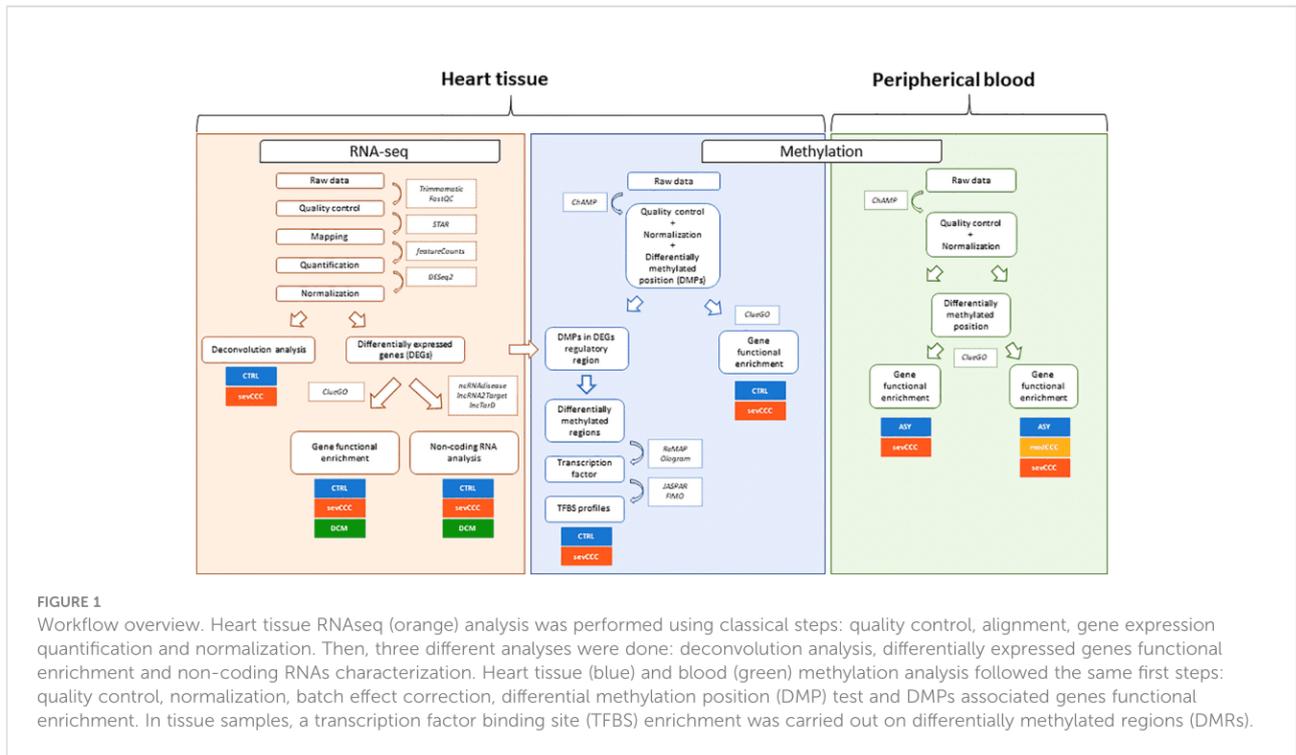


FIGURE 1 Workflow overview. Heart tissue RNAseq (orange) analysis was performed using classical steps: quality control, alignment, gene expression quantification and normalization. Then, three different analyses were done: deconvolution analysis, differentially expressed genes functional enrichment and non-coding RNAs characterization. Heart tissue (blue) and blood (green) methylation analysis followed the same first steps: quality control, normalization, batch effect correction, differential methylation position (DMP) test and DMPs associated genes functional enrichment. In tissue samples, a transcription factor binding site (TFBS) enrichment was carried out on differentially methylated regions (DMRs).

133a, RP11-276H19.1-GAS1, XIST-WNT1, MIR155HG-miR-155, LINC00707-ELAVL1 and KB-1732A1.1-E2F1). A similar analysis was performed on the ncRNAs associated to DCM (Supplementary Table 7 and Supplementary Figure 3B). 143/179 ncRNAs are specific to CCC (Supplementary Figure 3C).

DNA methylation directly affect the expression of some DEGs

Given the importance of methylation in the regulation of genetic response (18), we also performed several DNA

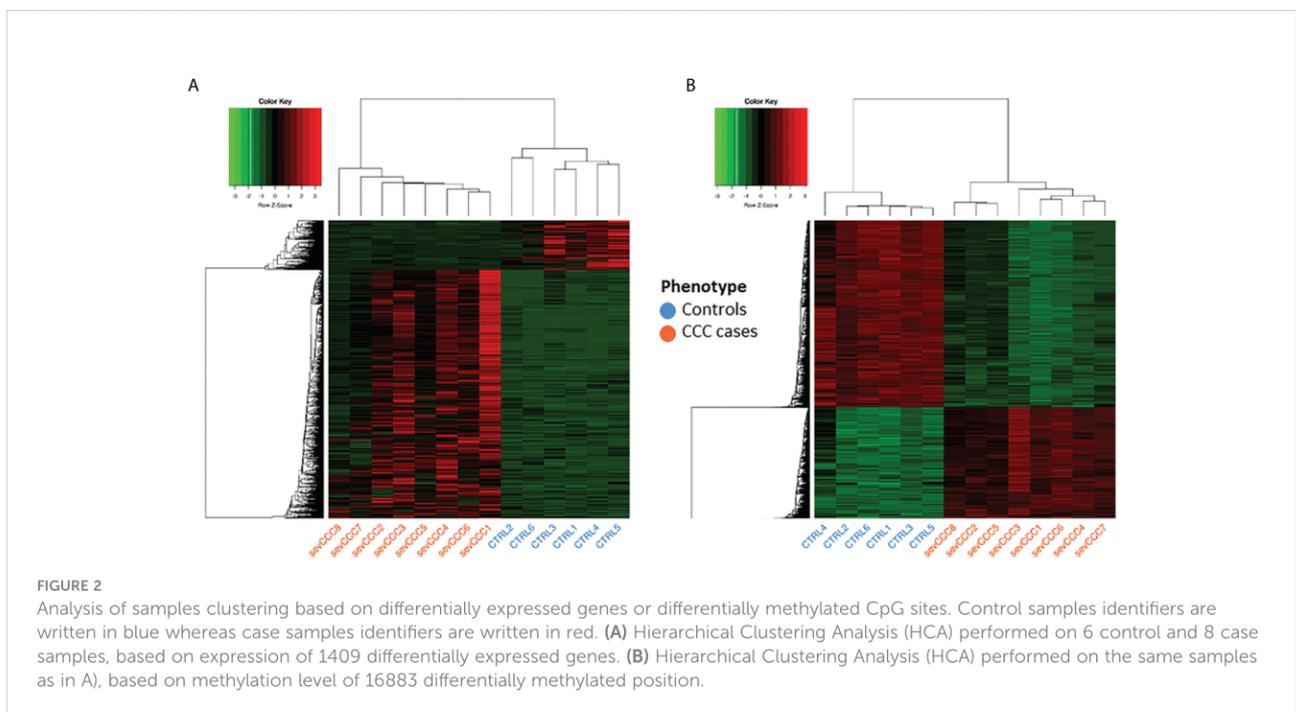
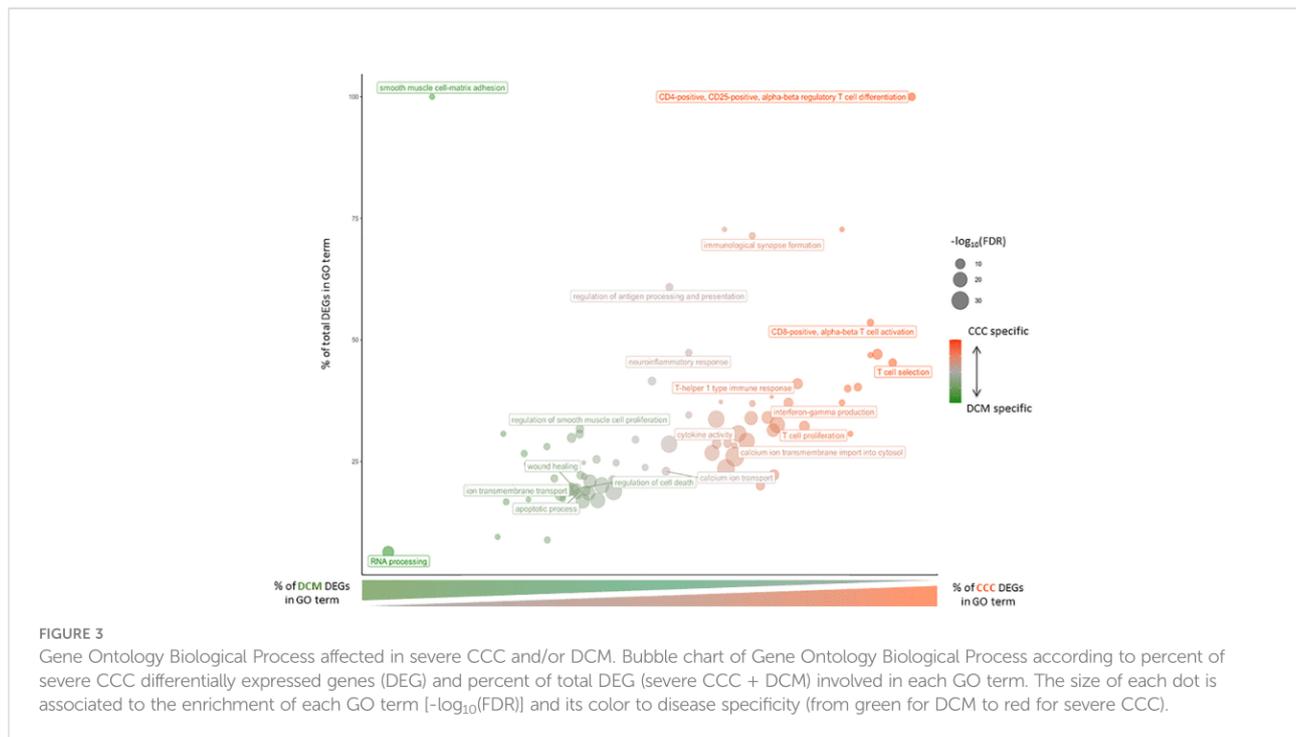


FIGURE 2 Analysis of samples clustering based on differentially expressed genes or differentially methylated CpG sites. Control samples identifiers are written in blue whereas case samples identifiers are written in red. (A) Hierarchical Clustering Analysis (HCA) performed on 6 control and 8 case samples, based on expression of 1409 differentially expressed genes. (B) Hierarchical Clustering Analysis (HCA) performed on the same samples as in A), based on methylation level of 16883 differentially methylated position.



methylation analyze in controls and CCC patients. Tissue sample DNA methylation analysis was conducted on the same samples as gene expression. Only 16883 CpGs (2.34%) were differentially methylated ($\text{FDR} < 0.05$ and $|\Delta\beta| > 0.2$) (Supplementary Table 8) and were associated to 5814 genes. PCA and HCA analysis confirmed the impact of those DMP on CCC (Figure 2B and Supplementary Figure 2D). The sex and the age of the patients seems to have no impact on this clustering (Supplementary Figures 2E, F).

Among the 16883 identified DMPs, 996 DMPs are associated to 390 DEGs. Interestingly, the upstream region of those DEGs (from TSS-1500 to TSS, 1st exon and 5'UTR) are significantly enriched in down-methylated DMPs ($\text{FDR} \leq 0.05$). Moreover, these down-methylated DMPs are, in 86% of the cases, associated with over-expressed DEGs. We retained these regions called “regulatory region” for further analysis.

Identification of transcription factors potentially involved in severe CCC

Based on the 409 DMPs associated with DEG regulatory regions, we were able to define 92 regulatory DMRs (Supplementary Table 9). These DMRs span on average 245bp and are in the promoter regions of 89 DEGs. Three analyses were performed with OLOGRAM to determine whether the enrichment in TF is significantly different of which would be expected by chance. Two analyses were performed: (A) with the promoters of the DEGs and (B) with the DMRs alone. For both analyses,

enrichment was compared with all genomic promoters. (Figure 4A). The $\log_2(\text{FC})$ s of each TF in analysis (A) and (B) are significantly correlated (Spearman p -value ≤ 0.05), but the r^2 is low (0.49) (Figure 4B) suggesting a different signal carried by the DMRs compared to all the promoters including the DMRs. Moreover, the $\log_2(\text{FC})$ s obtained in analysis B are significantly higher than the ones obtained in analysis A (Wilcoxon p -value = 4.23×10^{-7}) (Figure 4C). These two analyses clearly show that there is a stronger enrichment of TFs in DMRs than in the promoter set. To confirm these results, a third analysis (C) was performed with the DMRs compared to the promoters of the DEGs containing these DMRs. The obtained distribution of the $\log_2(\text{FC})$ with this analysis showed two distinct peaks, one around 0 and another around 1.7 (Figure 4C), confirming that some TFs are specifically enriched in DMRs. A total of 30 TFs were found significantly associated to the DMRs in the analyses B and C. We also considered as a Cis-Regulatory Module (CRMs) the regions where at least 2 TFs bind to the genome according to ReMap. After data filtering and considering combinations of up to 4 TFs, we have identified 16 regions significantly associated with our DMRs, involving a total of 12 transcription factors (Supplementary Table 10). The top-regulators are BRD4, EED, BCLAF1, TBX21, RUNX3 and RUNX1.

Immune, heart-relative or neurological processes affected by TFBS methylation

The 30 TF previously identified are involved in several biological process, such as somatic recombination of

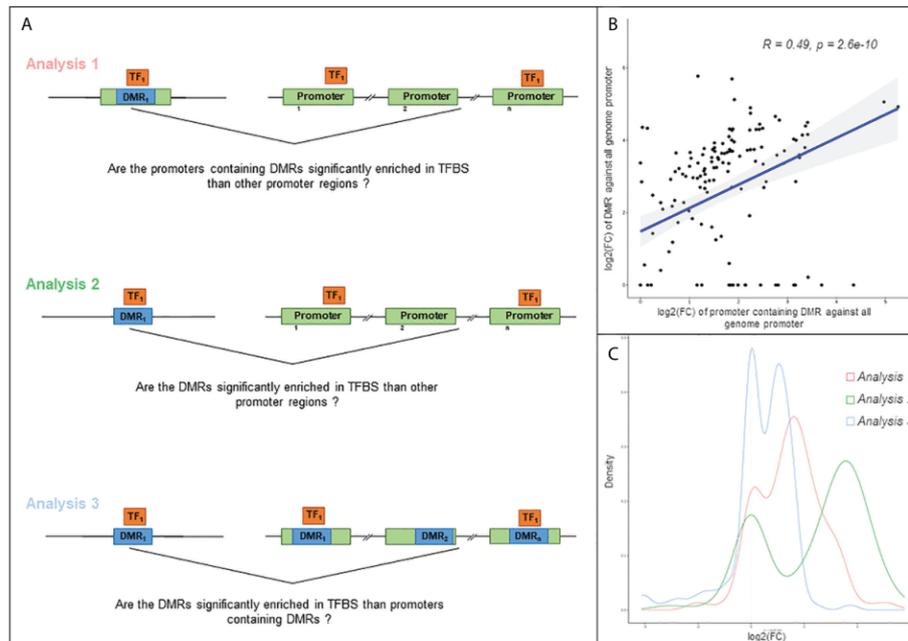


FIGURE 4

Analysis of the relation between TFBS (Transcription Factor Binding Site). (A) Schematic illustration of the three approaches used in this analysis. Differentially methylated region (DMR) is highlighted in blue, gene regulatory region in green, and transcription factor (TF) in orange. Analysis 1: TFBS enrichment in regulatory region containing at least one DMR, compared to all genome regulatory region. For each gene, a regulatory region is defined as the region from TSS-1500 to first exon. Analysis 2: TFBS enrichment in DMR compared to all genome regulatory region. Analysis 3: TFBS enrichment in DMR compared to regulatory region containing at least one DMR. (B) Scatter plot of the log₂(FC) obtained with the analysis 1 and 2 and Spearman correlation of these values. The fold change is computed according to the observed S value compared to obtained S value, S corresponding to the number of overlapping bases between TFBS and query region. (C) Distribution of the log₂(FC) obtained with the 3 approaches.

immunoglobulin gene segments, regulation of cardiac muscle tissue growth or peripheral nervous system neuron development (Supplementary Table 11). Among the genes involved in the Th1/IFN- γ response, 19 genes are potentially targeted by 28 of the 30 TF previously identified, illustrating the important differential methylation of immune response-related genes in the pathogenic process associated with CCC. Seven TFs (BCLAF1, BRD4, CBF, EED, PAX5, RUNX3 and TBX21) target at least 11/19 genes (Figure 5). A few of them are targeted by specific TFs (for example IFN- γ is targeted RUNX3+TBX21). Similar data were obtained for heart-relative or neurological process (Supplementary Figure 4).

5 TFBS identified in RUNX3 regulatory region

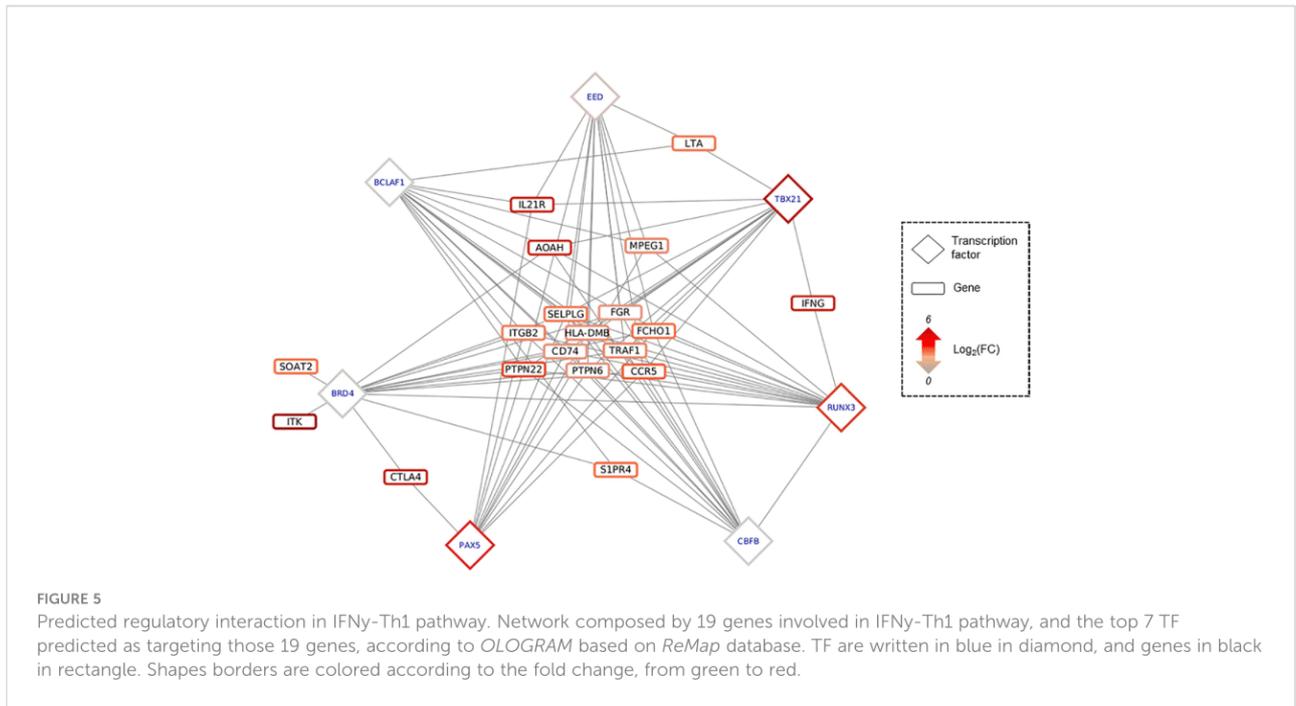
Of the 30 TFs of interest, 20 have at least one known motif in the Jaspar database, providing a total of 45 distinct motifs. After filtering predicted TFBS in our DMR overlapping at least one DMP, 423 TFBS have been identified, for all the 20 TFs (Supplementary Table 12). Interestingly, TBX21, RUNX3 and

EBF1 are the TFs whose binding motif appears to be affected by DNA methylation in the largest number of genes. Because these TFs are involved in many complexes, regulation of their binding may affect the binding of all TFs in the complexes.

Among the 25 genes with a low number of TFs binding in their promoter region ($n \leq 3$), and thus being affected by specific TFs, 7 are targeted by TBX21 and 4 by RUNX3, showing again the importance of these TFs in CCC. Considering that RUNX3, a key regulator in CCCs, also has a DMR in its promoter region, further analysis was performed on this TF. On this 831 bp DMR, 5 TFBSs are present (Figure 6). This DMR is targeted by at least 6 of the following 7 TFs: RUNX3, PAX5, YY1, SP1, MAX, EBF1 and IRF4. SP1 and PAX5 seem to be the most affine with these sequences. PAX5 is always found associated to IRF4, suggesting an interaction between those 2 TFs.

Immune cell type infiltration occurs in CCC heart tissue

Given the infiltration of immune cells in CCC myocardium, we characterized and quantified the proportions of these cell

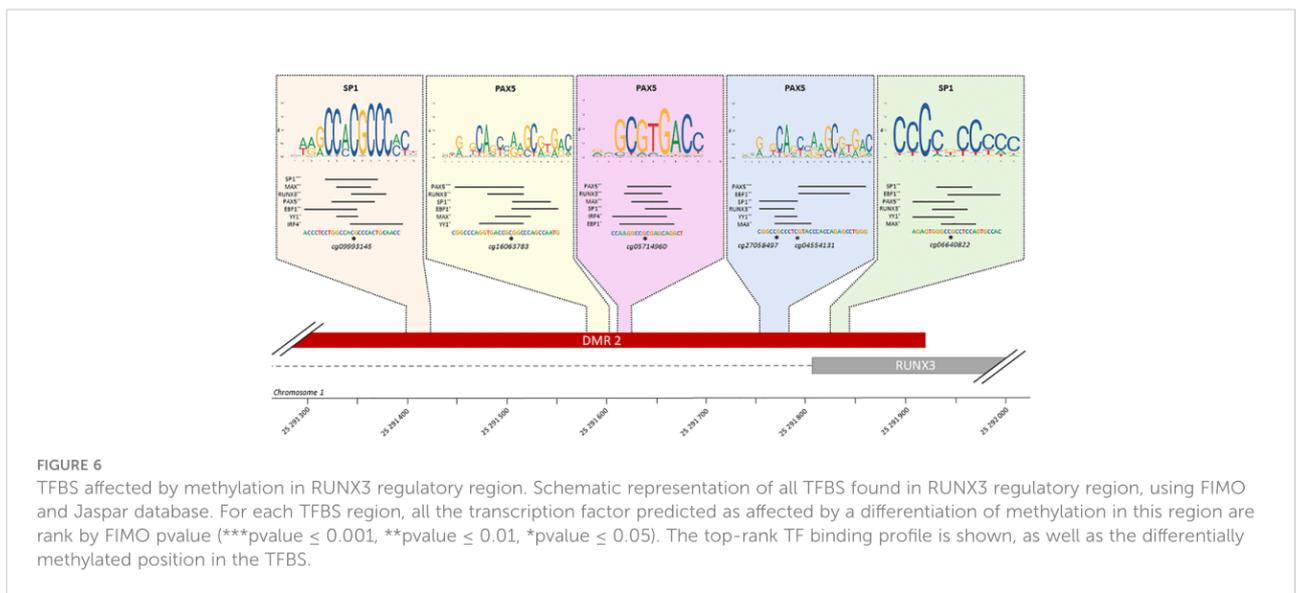


types in our samples. First of all, immune cell signatures (19) showed in general a higher proportion of immune cells in CCCs compared to controls (Supplementary Figure 5A), such as activated NKs, and more interestingly T CD8, T Cell memory and T follicular helper, with a reduced proportion of M2 macrophages. Secondly, a human heart tissue (left ventricle) cell signature (20) showed that the CCCs had fewer cardiomyocytes and smooth muscle than the controls (Supplementary Figure 5B). Moreover, CCC myocardium had a higher proportion of macrophages, in line with the immune

cell infiltration in the cardiac tissues on CCC patients (Supplementary Table 13).

Methylation sites in blood are associated to moderate or severe CCC

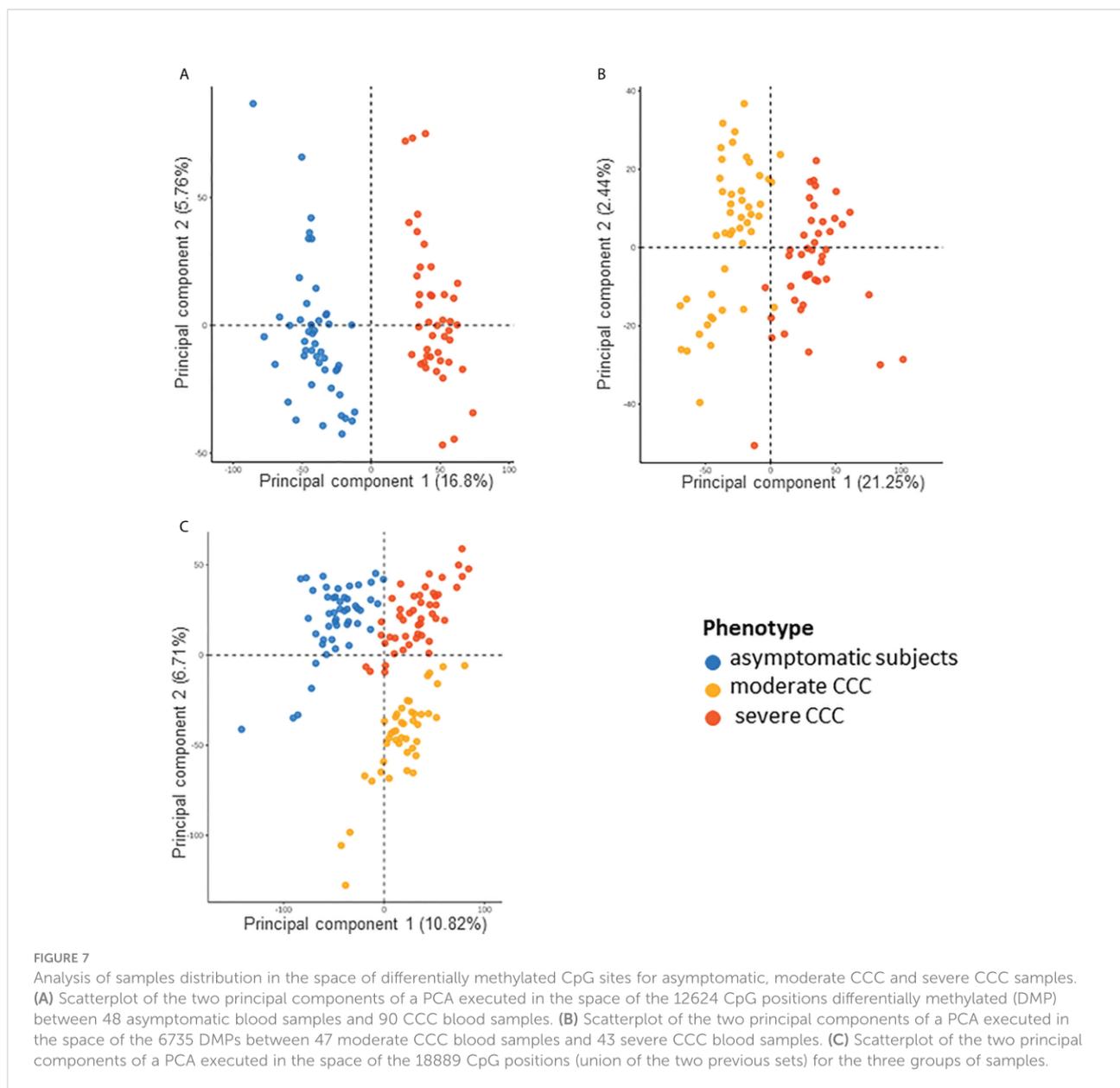
We also studied the DNA methylation in the blood of asymptomatic, moderate and severe CCC by hypothesizing that the blood data reflect the phenotype. We found 12624

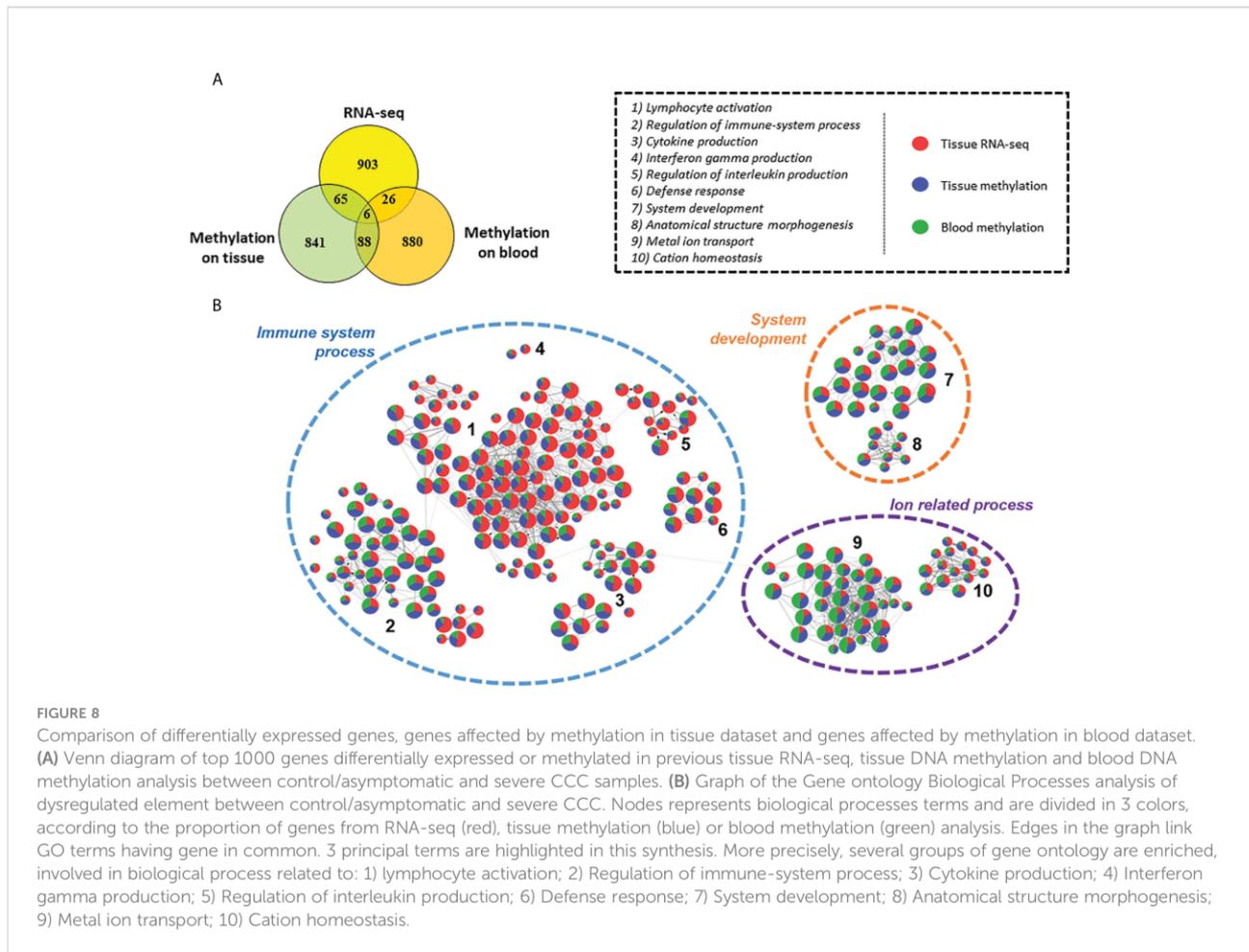


DMPs between asymptomatic and severe CCC blood samples ($FDR \leq 0.05$) (Supplementary Table 14). Despite the small variation in the level of DNA methylation detected in the blood ($\Delta\beta$), the methylation of these 12624 DMPs was enough to separate controls from cases, either *via* PCA (Figure 7A) or HCA (Supplementary Figure 6A). Association was found in 6436 genes with at least one DMP, but only 139 genes are in common between the three analyses (RNA-seq in tissue and DNA methylation in tissue and blood).

6735 CpGs were found as DMPs between moderate and severe CCC ($FDR < 0.05$) (Supplementary Table 15). They were

enough to discriminate samples according to the stage of the disease on a PCA (Figure 7B) or HCA (Supplementary Figure 6B). Only 470 DMPs were also identified as DMP in controls vs severe CCC blood methylation analysis, and 1750 genes out of 3911 (44.75%) are shared by both blood methylation analysis (non-significant enrichment). Interestingly, looking at the DNA methylation level of all DMPs found in blood (merge of the two previous analysis, $n=18889$) allowed us to distinguish the three groups of individuals, according to their phenotype (Figure 7C), revealing a gradient of methylation from controls to severe CCC through moderate CCC.





Nervous system related processes are strongly affected in moderate CCC

The general pattern detected on heart tissues and on blood are different and seemed to be tissue/fluid specific (Figure 8A). However, we conducted a Gene ontology analysis of the top 1000 genes of each analysis (Supplementary Table 16). 3 major biological processes are affected in our analyses: immune system, system development and ion transport (Figure 8B). They are shared by all three analyses, showing that although few genes are found in common, they are involved in common biological functions. Therefore, it seemed reasonable to analyze the methylation differences between moderate and severe CCC to understand the development of the disease (Supplementary Table 17). Unlike to the results found between controls and severe CCC, here genes are mostly involved in neurogenesis, cardiovascular system development or actin filament organization. These genes are associated with the immune response, notably in adaptive immune response and also with ion-related processes.

Discussion

This study focused on the study of epigenetic regulation in Chagas disease, and its impact on the development of the disease. Based on RNAseq and methylation data, 1409 differentially expressed genes (DEGs) were identified, among which 89 have differentially methylated regions (DMRs) in their regulatory region. 30 transcription factors were identified as potentially affected by these methylation differences, suggesting their involvement in genetic deregulation. Furthermore, similar biological processes are affected by differences in methylation in cardiac tissue or blood. Therefore, the pathogenic process in patients with moderate CCC could be studied.

In a previous analysis, combining microarray and methylation data, we showed that patients with CCC had significant differences in gene expression and DNA methylation compared to healthy controls (9). DEGs with DMPs were predominantly associated with the immune response, but also with several biological processes such as arrhythmia, muscle contraction, fibrosis and mitochondrial function. Here, we have set up a more advanced analysis,

allowing us to have an in-depth analysis of genomic dysregulation. To do so, we were interested in both DNA methylation in order to identify transcription factors involved in the pathogenic process and in non-coding RNAs. Indeed, non-coding RNA, notably lncRNA and miRNA, are known to be involved in many of cardiovascular diseases (21, 22). The presented results extend the list of miRNAs associated to CCC, and not differentially expressed in DCM (10, 21, 23–27). Among the 19 miRNAs identified in this study, 3 have already been associated with severe CCC: miR-223, miR-208 and miR-151 (6). The involvement of the long non-coding RNA MIAT, previously associated to CCC (11), has also been confirmed. It acts as a sponge of miRNA-133 in breast cancer, dysregulated in CCC (10), and down-regulated in patients with heart failure (28).

Three distinct biological processes were associated to CCC, either we look at gene dysregulation or methylation, in heart tissue or blood: immune response, ion transport and nervous system. Although ion transport and nervous system were also partially associated to DCM, immune response remains specific to CCC. Moreover, the genes affected in those pathways were not the same between these diseases, illustrating a specific response in CCC. For example, several genes associated to ion transport, and more precisely with potassium voltage-gated ion channel, like KCNA2, KCNA5, KCNAB2, KCNB2, KCNC2, KCNG3 and KCNN4, were found to be differentially expressed and/or methylated in CCC, and not in DCM. Of interest, while KCNB2, KCNC2 and KCNG3 are mainly expressed in neurons, KCNA5 is expressed in heart muscle and dendritic cells, while KCNA2, KCNAB2 are expressed in neurons and immune cells, indicating that potassium channels of neuronal, cardiomyocyte and inflammatory cells are modulated in CCC heart tissue. Dysregulation of potassium channel genes were reported in heart of 36 *T. cruzi* infected mice (21) and 25 severe CCC heart (9). Dampening of calcium and potassium ion channels in mouse heart tissue led to a reduced Ca²⁺ release and prolongation of action potential (29). TNF- α may amplify this dysfunction by inducing nitric oxide synthase (NOS2) and oxidant species, which promotes electrophysiological changes in rat ventricular myocyte (30). KCNN4 (KCa3.1), differentially expressed in CCC (not in DCM) and containing DMP in its promoter, acts as a regulator of membrane potential in T cells. After antigen recognition by the T cell repertoire, the Ca²⁺ enters in the cytosol and sequentially activate KCa3.1 and Kv1.3 (KCNA3), causing membrane depolarization (31). This calcium, present in cytosol, will activate the NFAT protein, which regulate genes involving in T cell activation (32). Moreover, high levels of Kv1.3 were found in multiple autoimmune diseases, such as multiple sclerosis or diabetes (33–35). In our data, these two potassium channels are up-regulated, and KCa3.1 seems to be targeted by RUNX3. The potential biomarker for heart disease Galectin-3 (Gal-3), upregulating

the expression and activity of KCa3.1 channel in inflammatory cells and fibroblasts. Upregulated KCa3.1 facilitates inflammatory cell infiltration into the myocardium and fibroblast differentiation into activated form. Recently, it was shown that mice treated with TRAM-34, a KCa3.1 channel-specific inhibitor, either for 1- or 2-month period effectively reduced collagen deposition (36). Moreover, Gang et al. had shown that KCa3.1 inhibition by TRAM-34 therapy attenuated the increased inflammatory cell infiltration. Rezende de Castro et al. (37) studied gene expression profiling in acutely *T. cruzi* infected mice (15-day post-infection) and have found enriched pathways related to immunity and Th1 T cell immunity, cardiac conduction including potassium channels, protein synthesis and mitochondrial genes. Even though acute infection by *T. cruzi* (where parasites are very numerous in the myocardium) is not a perfect match for CCC, where parasites are scarce or absent from heart tissue, modulation of inflammatory cell and potassium channel gene expression profiles are similar in both situations.

On the other hand, our study underlined that several genes associated with moderate or severe CCC are involved in the nervous system. This enrichment is even stronger in moderate CCCs. Actually, the loss of neuron cells is a well-known phenomenon in digestive forms of Chagas disease (38, 39). In human hearts with dilated cardiomyopathy, the number of neuron cells is significantly reduced compared to controls. This neuronal depopulation is even more important in CCC (40), and during the acute phase in the mouse model (41). Similarly, a study showed that acute phase *T. cruzi* infected dogs displayed a rarefaction of the noradrenergic and acetylcholine nerve terminals accompanied by myocarditis, and the ventricles sympathetic denervation was present when the inflammatory process was moderate to intense (42). Moreover, in the rat model, moderate myocarditis lasting for two weeks caused complete denervation (43). A more recent study demonstrated that knockdown of acetylcholine in mice increased the Th1 response and clearance of *T. cruzi* parasitism in blood and tissue, but worsens the cardiac lesions and inflammatory infiltration (44). In the continuity of these observations, the increase in the amount of acetylcholine was associated with reduction of heart weight, inflammatory infiltration, and the fibrosis area, suggesting a reversing of cardiac hypertrophy (44). Moreover, macrophages, Th1, CD4 and B cells express adrenergic receptors (45–47). Their stimulation induces an increase of cAMP, which inhibits NF- κ B activation, leading to the suppression of type 1 pro-inflammatory cytokines and promoting the production of type 2 anti-inflammatory cytokines, such as IL10. This cytokine inhibits the antigen presenting capacity of macrophages and dendritic cells, and thus the differentiation of CD4 T cells to Th1 (48). Finally, circulating antibodies binding acetylcholine and norepinephrine receptors have been found in Chagas patient serum, suggesting an autoimmune response (49), and correlated with the previous study, inducing the lack of neurotransmitters,

and then the over production of Th1 cells. In our data, the genes related to acetylcholine production, or acetylcholine transporter are not differentially methylated or expressed in tissue samples according to fold change cut-off but present an $FDR < 0.05$. In blood methylation samples, ACHE (acetylcholine production), and CHRNA3, CHRNA4 and CHRNA7 (acetylcholine receptors) genes have differentially methylated positions in the promoter or body. While CHRNA4 is mainly expressed in neurons, ACHE is found in neurons and cardiac/skeletal muscle, CHRNA7 is mainly expressed in neurons, cardiac/skeletal muscle and immune cells, and CHRNA3 is found in neurons, and thymus.

Both potassium voltage-gated ion channel and nervous system identified in the current and previous studies seem to be linked to immune response, and more precisely T-cell related process. Moreover, in our data, the immune response was strongly stimulated, especially the activation of T cells even in the absence of the parasite. This inflammatory infiltrate was already reported elsewhere (50) and is mainly composed of Th1 lymphocytes, macrophages, and NK-cells. In this work, we performed a deconvolution analysis on our heart tissue collection, and we confirmed the content of this infiltrate. Our group, among others, had already shown that the T-cell infiltrating heart tissue strongly produced IFN- γ and TNF- α (4). In parallel, lower quantities of IL-2, IL-4, IL-6 and IL-10 were detected in CCC heart tissue (51). Our results are consistent with current knowledge of the pathology, with high expression of genes involved in IFN- γ and cytokine production. According to Gene Ontology Biological Process enrichment, genes affected in CCC are also involved in interleukin production, including IL-2, IL-4 and IL-6. IFN- γ , which has been described simultaneously as a pathogen resistance and a disease tolerance gene is also acting as an upstream regulator (52). Indeed, IFN- γ stimulates the inflammatory response indirectly *via* the NF- κ B pathway and activate the production of ROS and NOS, which, in an excessive quantity, have also deleterious effect on mitochondria and cardiomyocytes (5). Interestingly, IFN- γ , with IFN- β and TNF- α (53), induced expression of miR-155 (53). In our data, MIR155HG, coding for this miRNA, is up-regulated. miR-155 is also up-regulated in viral myocarditis; where it is expressed by infiltrating immune cells, and seems to be involved in TNF- α , IFN- γ and IL-6 production, as well as immune cell infiltration (54). The lack of this microRNA seems to decrease IFN- γ and TNF- α in the acute stage of Chagas cardiomyopathy in mice heart tissue (21). Moreover, we found a DMR in IFN- γ promoter region, targeted TBX21. According to Ologram, RUNX3 could also fix this DMR, but its TFBS wasn't found in this DMR, as described in JASPAR database. Those two transcriptions factors are involved in Th1 differentiation, as well as GATA3 (55, 56). According to our analysis, the micro-RNA miR-142, associated to Th1 differentiation (57) in neuronal autoimmune disease,

potentially targets RUNX3 and TBX21. In Chagas disease, TBX21 and IFN- γ expression are correlated with the left ventricular dilation, and the ratio between TBX21 and GATA3 expression is significantly higher in CCC than in non-inflammatory cardiomyopathy (58), which were also confirmed in our data in both control to CCC or DCM comparisons. Moreover, RUNX3 overexpression has been associated to the methylation of its promoter in CCC (9), as well as in our current analysis. In total, RUNX3 and TBX21 targets 29 genes, including 14 in common (AIM2, ARHGAP30, BATF, C16orf54, CCR5, CYTIP, DENND2D, DOK2, FGR, MIR142, PTPN22, TRAF1, TRAF3IP3 and WDFY4). Moreover, RUNX3 is targeted by 7 transcription factors: EBF1, IRF4, MAX, PAX5, RUNX3, SP1 and YY1.

We also found that XIST is one of the tops up-regulated lncRNA in Chagas disease. It is associated with the WNT1 gene, belonging to the Wnt family, but because of the wide sex dispersion in our dataset, we cannot conclude the direct involvement of this lncRNA in CCC. In our data, WNT1, as well as WNT10A and WNT10B are up-regulated in severe CCC. The Wnt pathway is involved in the differentiation of Th2 cells (59) and its inhibition in acute Chagas disease stage decrease the Th2 response (60) and increase inflammatory response, controlling the parasite (61). Studies have demonstrated that Wnt activation is related to pathological stages including inflammation, angiogenesis, and fibrosis and aberrant expression is associated with cardiovascular diseases (61, 62). Indeed, Analysis of the expression of Wnt proteins indicated that Wnt-2, Wnt-4, Wnt-10b, and Wnt-11 were significantly upregulated 5 days after myocardial infarction (63). Studies confirmed the interactions between the Wnt pathway and TGF- β signalling (64–66). TGF- β receptor overexpression has already been observed in acute phase of CCC (67). It has several role in CCC development, including parasitic invasion, inflammation, immune response, heart fibrosis and heart conduction (68). Our results had also shown the potential implication of NOTCH1 (downregulated and targeting DEGs with DMRs in their regulatory regions) in the genetic dysregulation in severe CCC. Activation of Notch signalling limits the range of cardiac damage by the improving of angiogenesis (69–71), cardiac regeneration and cardio protection (72), and reducing fibrosis (73), apoptosis (74), and oxidative stress (75).

Few studies have linked Notch signaling to immune response and parasitic infections. Tu et al. has shown that in *Trichuris muris* infection, a deficiency of Notch signaling in T cells led to a failure in initiating the Th2 response (76). It suggests that Notch signaling influences the Th2 profile and alters immune responses against parasites. Similarly, during Leishmania major infection, the Th1 response was induced in deficient mice, which presented high levels of IFN- γ that led to infection control (76). So, we can raise the hypothesis that a

down expression of NOTCH1 in CCC patients promote the Th1 response associated to severe cardiomyopathy. To support this hypothesis, Guzmán-Rivera et al. had treated Chagas infected mice with simvastatin (77). Simvastatin activates the Notch 1 pathway in the hearts of *T. cruzi* infected mice and decreases the cellular infiltrate, inflammatory cytokines and prevents the increase in collagen deposition in cardiac tissue.

All the results obtained in our study converge towards a combined involvement of processes related to the immune response, ion transport, cardiac contraction and the nervous system. In particular, the nervous system appears to be strongly impacted between moderate and severe CCC, and the potassium-related process specific to CCC compared to DCM. Gene expression analysis alone has revealed some non-coding elements but has not provided so much new information about the disease. The inclusion of methylation put forward less obvious biological processes. Thus, most of the identified genes (differentially expressed and/or methylated) are generally involved in several of these processes, highlighting links between them. More precisely, all the previously described biological process seems to be linked to the immune response, and notably to the Th1 response, including IFN- γ and RUNX3. These results, combined with those obtained in previous analyses of CCC (9, 78), confirm the importance of DNA methylation in the development of CCC.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/[Supplementary Material](#).

Ethics statement

The protocol was approved by the institutional review boards of the University of São Paulo School of Medicine and INSERM (French National Institute of Health and Medical Research). Written informed consent was obtained from all patients. All experimental methods comply with the Helsinki Declaration.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication. Study design: PB, JK, LS, ECN, CC. Phenotype characterization: BMI, CM, RHBS, SS, ANS, PP, AIF, EAB, CWP, BS, FCD, MFS, FAG, JAMN, AF, SS, GDLP, FB, PB, HTLW, AS, MM, MHH, EAD, ACP, VRJ, Experimental

analysis: PB, LL, AFF, JPSN, PCT, LRPF, AK, SC, DDSC, RCFZ, VOCC, RRA. Statistical analysis: PB, QF, DP, LS, ECN, CC. Manuscript preparation: PB, LS, ECN, CC.

Funding

This work was supported by the Institut National de la Santé et de la Recherche Médicale (INSERM); the Aix-Marseille University (grant number: AMIDEX “International_2018” MITOMUTCHAGAS); the French Agency for Research (Agence Nationale de la Recherche-ANR (grant numbers: “Br-Fr-Chagas”, “landscardio”); the CNPq (Brazilian Council for Scientific and Technological Development); and the FAPESP (São Paulo State Research Funding Agency Brazil (grant numbers: 2013/50302-3, 2014/50890-5); the National Institutes of Health/USA (grant numbers: 2 P50 AI098461-02 and 2U19AI098461-06). This work was founded by the Inserm Cross-Cutting Project GOLD. This project has received funding from the Excellence Initiative of Aix-Marseille University - A*Midex a French “Investissements d’Avenir programme”- Institute MarMaRa AMX-19-IET-007. JPSN was a recipient of a MarMaRa fellowship. EC-N and JK, are recipients of productivity awards by CNPq. The funders did not play any role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgments

Center de Calcul Intensif d’Aix-Marseille is acknowledged for granting access to its high performance computing resources. Authors thanks Delphine Potier, Pierre Milpied, Benoit Ballester for helpful discussions and comments.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2022.958200/full#supplementary-material>

SUPPLEMENTARY TABLE 1

Biological samples included in this study.

SUPPLEMENTARY TABLE 2

Cell lines description.

SUPPLEMENTARY TABLE 3

List of unique differentially expressed genes identified by RNA-seq between 8 end stage CCC heart tissue samples and 6 heart tissue samples obtained from organ donors.

SUPPLEMENTARY TABLE 4

Statistics on coding sequences detected by RNA-seq and on differentially expressed genes.

SUPPLEMENTARY TABLE 5

List of unique differentially expressed genes identified by RNA-seq between 8 DCM heart tissue samples and 6 heart tissue samples obtained from organ donors.

SUPPLEMENTARY TABLE 6

Gene ontology analysis based on differentially expressed genes between control and severe CCC and/or between control and DCM.

SUPPLEMENTARY TABLE 7

List of differentially expressed ncRNAs between control and severe CCC, or between control and dilated cardiomyopathies.

SUPPLEMENTARY TABLE 8

List of the Differentially Methylated CpGs on heart tissues samples (controls versus CCC).

SUPPLEMENTARY TABLE 9

List of the regulatory DMRs associated to the 89 DEGs.

SUPPLEMENTARY TABLE 10

List of transcription factor and transcription factor complexes testing against our DMRs.

SUPPLEMENTARY TABLE 11

Gene ontology analysis of the TFs binding our DMRs.

SUPPLEMENTARY TABLE 12

For each DMR, list of predicted Transcription Factor Binding Site (TFBS) and their associated Transcription Factor (TF).

SUPPLEMENTARY TABLE 13

Cell type proportions in RNA-seq by comparison to cell type signatures.

SUPPLEMENTARY TABLE 14

List of the Differentially Methylated CpGs between blood asymptomatic samples and blood CCC samples.

SUPPLEMENTARY TABLE 15

List of the Differentially Methylated CpGs between blood moderate CCC samples and blood severe CCC samples.

SUPPLEMENTARY TABLE 16

Gene ontology analysis based on differentially expressed genes between control and severe CCC and/or genes affected by at least one differentially methylated CpGs between tissue control and severe CCC and/or genes affected by at least one differentially methylated CpGs between blood asymptomatic and severe CCC samples.

SUPPLEMENTARY TABLE 17

Gene ontology analysis based on genes affected by at least one differentially methylated CpGs between blood moderate CCC samples and blood severe CCC samples.

SUPPLEMENTARY FIGURE 1

Volcanoplots of differential expression testing for CCC and DCM. Volcanoplot showing the difference of gene expression between (A) controls and CCCs or (B) controls and DCMs. On the x-axis is represented the $\log_2(\text{FC})$ and on the y-axis the $-\log_{10}(\text{adjusted pvalue})$. Overexpressed DEGs are in red, and underexpressed in green.

SUPPLEMENTARY FIGURE 2

Analysis of samples distribution in the space of differentially expressed genes or differentially methylated CpG sites between control and severe CCC. Age and sex impact on gene expression or on methylation. (A) Scatterplot of the two first principal component of a PCA of 6 Control and 8 Case samples executed in the space of the 1409 differentially expressed genes between the two conditions. (B) Scatterplot of the two first principal component of a PCA of 6 Control and 8 Case samples executed in the space of the 1409 differentially expressed genes between the two conditions colored by sex (blue: male, pink: female). (C) Scatterplot of the two first principal component of a PCA of 6 Control and 8 Case samples executed in the space of the 1409 differentially expressed genes between the two conditions colored by age (from grey to blue). (D) Scatterplot of the two first principal component of a PCA of 6 Control and 8 Case samples executed in the space of the 16883 differentially methylated CpG sites between the two conditions. (E) Scatterplot of the two first principal component of a PCA of 6 Control and 8 Case samples executed in the space of the 16883 differentially methylated CpG sites between the two conditions colored by sex (blue: male, pink: female). (F) Scatterplot of the two first principal component of a PCA of 6 Control and 8 Case samples executed in the space of the 16883 differentially methylated CpG sites between the two conditions colored by age (from grey to blue).

SUPPLEMENTARY FIGURE 3

ncRNA analysis in severe CCC and DCM. Control samples identifiers are written in blue, severe CCC samples identifiers in red and DCM samples in green. (A) Hierarchical Clustering Analysis (HCA) performed on 6 control and 8 severe CCC samples, based on expression of 179 differentially expressed ncRNAs. (B) Hierarchical Clustering Analysis (HCA) performed on 6 control and 8 DCM samples, based on expression of 327 differentially expressed ncRNAs. (C) Venn diagram of shared differentially expressed ncRNAs between severe CCC and DCM.

SUPPLEMENTARY FIGURE 4

Predicted regulatory interaction in cardiac muscular or nervous system process. Network composed by 6 transcription factors (TF) involved in cardiac muscular and/or nervous system process, and their 68 targeted genes, predicted by *OLOGRAM* according to *ReMap* database. TF are written in blue in diamond, and genes in black in rectangle. Shapes borders are colored according to the fold change, from green to red.

SUPPLEMENTARY FIGURE 5

Estimation of cell proportion in control and severe CCC heart tissue. Deconvolution of the RNA-seq bulk gene expressions of 6 control and 8 severe CCC samples to infer the proportion of cells in the samples. The * represent cell types whose proportion is significantly different (Wilcoxon test, $\text{FDR} < 0.05$) between controls and cases. (A) Deconvolution using cell type signature with 22 immunological cell types (LM22 signature matrix). (B) Deconvolution using cell type signature with 5 left ventricle related cell types.

SUPPLEMENTARY FIGURE 6

Analysis of samples clustering based on differentially methylated CpG sites in blood samples. Asymptomatic samples identifiers are written in blue, moderate CCC in orange and severe CCC in red. (A) Hierarchical Clustering Analysis (HCA) performed on 48 asymptomatic and 90 severe CCC samples, based on expression of 12624 differentially methylated position. (B) Hierarchical Clustering Analysis (HCA) performed on 47 moderate CCC and 43 severe CCC samples, based on methylation level of 6735 differentially methylated position.

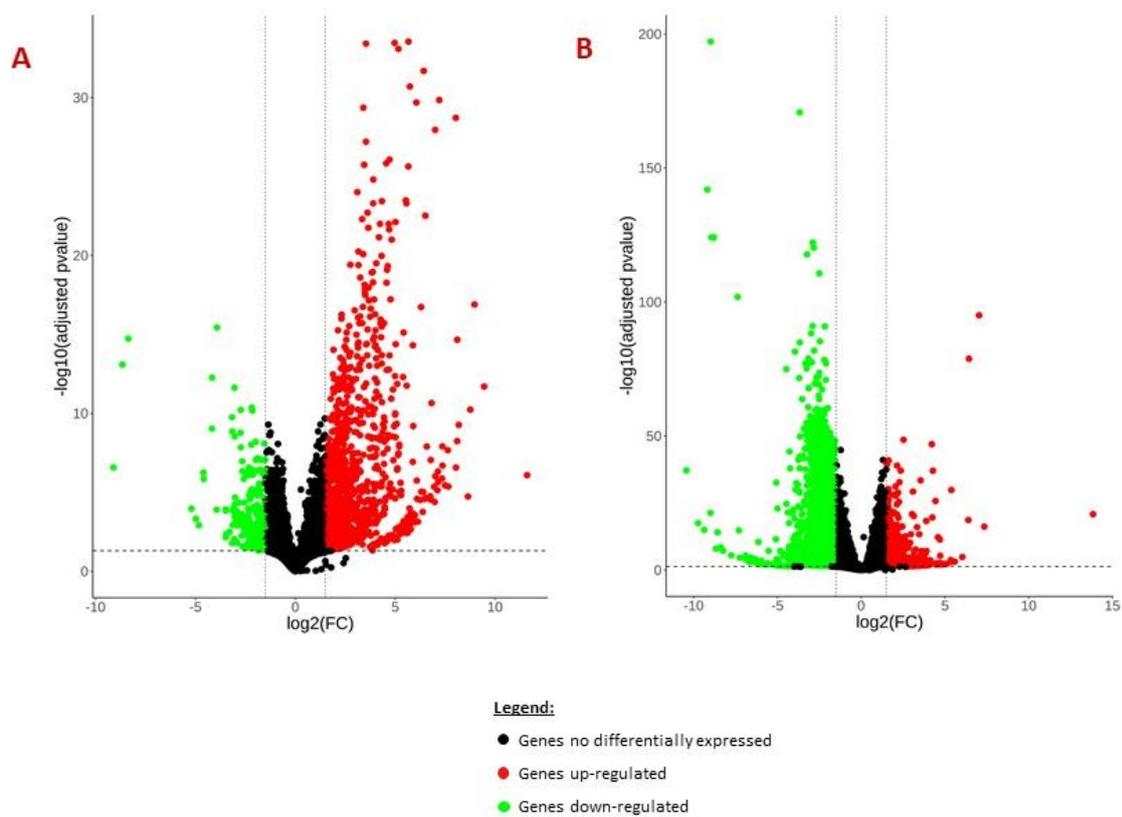
References

- Pérez-Molina JA, Molina I. Chagas disease. *Lancet* (2018) 391(10115):82–94. doi: 10.1016/S0140-6736(17)31612-4
- Freitas HFG, Chizzola PR, Paes ÂT, Lima ACP, Mansur AJ. Risk stratification in a Brazilian hospital-based cohort of 1220 outpatients with heart failure: role of chagas' heart disease. *Int J Cardiol* (2005) 102(2):239–47. doi: 10.1016/j.ijcard.2004.05.025
- Morillo CA, Marin-Neto JA, Avezum A, Sosa-Estani S, Rassi A, Rosas F, et al. Randomized trial of benznidazole for chronic chagas' cardiomyopathy. *New Engl J Med* (2015) 373(14):1295–306. doi: 10.1056/NEJMoa1507574
- Abel LC, Rizzo LV, Ianni B, Albuquerque F, Bacal F, Carrara D, et al. Chronic chagas' disease cardiomyopathy patients display an increased IFN- γ response to trypanosoma cruzi infection. *J Autoimmun* (2001) 17(1):99–107. doi: 10.1006/jaut.2001.0523
- Cunha-Neto E, Dzau VJ, Allen PD, Stamatou D, Benvenuti L, Higuchi ML, et al. Cardiac gene expression profiling provides evidence for cytokinopathy as a molecular mechanism in chagas' disease cardiomyopathy. *Am J Pathol* (2005) 167(2):305–13. doi: 10.1016/S0002-9440(10)62976-8
- Laugier L, Ferreira LRP, Ferreira FM, Cabantous S, Frade AF, Nunes JP, et al. miRNAs may play a major role in the control of gene expression in key pathobiological processes in chagas disease cardiomyopathy. *PLoS Negl Trop Dis* (2020) 14(12):e0008889. doi: 10.1371/journal.pntd.0008889
- Jo BS, Koh IU, Bae JB, Yu HY, Jeon ES, Lee HY, et al. Methylome analysis reveals alterations in DNA methylation in the regulatory regions of left ventricle development genes in human dilated cardiomyopathy. *Genomics* (2016) 108(2):84–92. doi: 10.1016/j.ygeno.2016.07.001
- Haas J, Frese KS, Park YJ, Keller A, Vogel B, Lindroth AM, et al. Alterations in cardiac DNA methylation in human dilated cardiomyopathy. *EMBO Mol Med* (2013) 5(3):413–29. doi: 10.1002/emmm.201201553
- Laugier L, Frade AF, Ferreira FM, Baron MA, Teixeira PC, Cabantous S, et al. Whole-genome cardiac DNA methylation fingerprint and gene expression analysis provide new insights in the pathogenesis of chronic chagas disease cardiomyopathy. *Clin Infect Dis* (2017) 65(7):1103–11. doi: 10.1093/cid/cix506
- Ferreira LRP, Frade AF, Santos RHB, Teixeira PC, Baron MA, Navarro IC, et al. MicroRNAs miR-1, miR-133a, miR-133b, miR-208a and miR-208b are dysregulated in chronic chagas disease cardiomyopathy. *Int J Cardiol* (2014) 175(3):409–17. doi: 10.1016/j.ijcard.2014.05.019
- Frade AF, Laugier L, Ferreira LRP, Baron MA, Benvenuti LA, Teixeira PC, et al. Myocardial infarction-associated transcript, a long noncoding RNA, is overexpressed during dilated cardiomyopathy due to chronic chagas disease. *J Infect Dis* (2016) 214(1):161–5. doi: 10.1093/infdis/jiw095
- Frade AF, Teixeira PC, Ianni BM, Pissetti CW, Saba B, Wang LHT, et al. Polymorphism in the alpha cardiac muscle actin 1 gene is associated to susceptibility to chronic inflammatory cardiomyopathy. *PLoS One* (2013) 8(12):e83446. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3868584/>. doi: 10.1371/journal.pone.0083446
- Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* (2014) 15(12):550. doi: 10.1186/s13059-014-0550-8
- Luo W, Friedman MS, Shedden K, Hankenson KD, Woolf PJ. GAGE: generally applicable gene set enrichment for pathway analysis. *BMC Bioinf* (2009) 10(1):161. doi: 10.1186/1471-2105-10-161
- Chêneby J, Gheorghe M, Artufel M, Mathelier A, Ballester B. ReMap 2018: an updated atlas of regulatory regions from an integrative analysis of DNA-binding ChIP-seq experiments. *Nucleic Acids Res* (2018) 46(D1):D267–75. doi: 10.1093/nar/gkx1092
- Ferré Q, Charbonnier G, Sadouni N, Lopez F, Kermezli Y, Spicuglia S, et al. OLOGRAM: Determining significance of total overlap length between genomic regions sets. *Bioinformatics* (2019), 36(6):btz810. doi: 10.1093/bioinformatics/btz810
- Ferré Q, Capponi C, Puthier D. OLOGRAM-MODL: mining enriched n-wise combinations of genomic features with Monte Carlo and dictionary learning. *NAR Genomics Bioinf* (2021) 3(4):lqab114. doi: 10.1093/nargab/lqab114
- Razin A, Cedar H. DNA Methylation and gene expression. *Microbiol Rev* (1991) 55(3):451–8. doi: 10.1128/mr.55.3.451-458.1991
- Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods* (2015) 12(5):453–7. doi: 10.1038/nmeth.3337
- Wang L, Yu P, Zhou B, Song J, Li Z, Zhang M, et al. Single-cell reconstruction of the adult human heart during heart failure and recovery reveals the cellular landscape underlying cardiac function. *Nat Cell Biol* (2020) 22(1):108–19. doi: 10.1038/s41556-019-0446-7
- Navarro IC, Ferreira FM, Nakaya HI, Baron MA, Vilar-Pereira G, Pereira IR, et al. MicroRNA transcriptome profiling in heart of trypanosoma cruzi-infected mice: Parasitological and cardiological outcomes. *PLoS Negl Trop Dis* (2015) 9(6):e0003828. doi: 10.1371/journal.pntd.0003828
- Hobuß L, Bär C, Thum T. Long non-coding RNAs: At the heart of cardiac dysfunction? *Front Physiol* (2019) 10:30/full. doi: 10.3389/fphys.2019.00030/full
- Ballinas-Verdugo MA, Jiménez-Ortega RF, Martínez-Martínez E, Rivas N, Contreras-López EA, Carbó R, et al. Circulating miR-146a as a possible candidate biomarker in the indeterminate phase of chagas disease. *Biol Res* (2021) 54(1):21. doi: 10.1186/s40659-021-00345-3
- Ferreira LRP, Ferreira FM, Laugier L, Cabantous S, Navarro IC, da Silva Cândido D, et al. Integration of miRNA and gene expression profiles suggest a role for miRNAs in the pathobiological processes of acute trypanosoma cruzi infection. *Sci Rep* (2017) 7(1):17990. doi: 10.1038/s41598-017-18080-9
- Jha BK, Varikuti S, Seidler GR, Volpedo G, Satoskar AR, McGwire BS. MicroRNA-155 deficiency exacerbates trypanosoma cruzi infection. *Infection Immun* (2020) 88(7):e00948-19. doi: 10.1128/IAI.00948-19
- Nonaka CKV, Macêdo CT, Cavalcante BRR, de Alcântara AC, DN S, Bezerra M da R, et al. Circulating miRNAs as potential biomarkers associated with cardiac remodeling and fibrosis in chagas disease cardiomyopathy. *Int J Mol Sci* (2019) 20(16):E4064. doi: 10.3390/ijms20164064
- Nonaka CKV, Sampaio GL, Silva KN, Khouri R, Macedo CT. Chagas Translational Research Consortium null, et al. Therapeutic miR-21 silencing reduces cardiac fibrosis and modulates inflammatory response in chronic chagas disease. *Int J Mol Sci* (2021) 22(7):3307. doi: 10.3390/ijms22073307
- Vogel B, Keller A, Frese KS, Leidinger P, Sedaghat-Hamedani F, Kayvanpour E, et al. Multivariate miRNA signatures as biomarkers for non-ischaemic systolic heart failure. *Eur Heart J* (2013) 34(36):2812–22. doi: 10.1093/eurheartj/ehz256
- Roman-Campos D, Sales-Júnior P, Duarte HL, Gomes ER, Guatimosim S, Ropert C, et al. Cardiomyocyte dysfunction during the chronic phase of chagas disease. *Mem Inst Oswaldo Cruz* (2013) 108(2):243–5. doi: 10.1590/0074-0276108022013019
- Fernández-Velasco M, Ruiz-Hurtado G, Hurtado O, Moro MA, Delgado C. TNF- α downregulates transient outward potassium current in rat ventricular myocytes through iNOS overexpression and oxidant species generation. *Am J Physiol Heart Circ Physiol* (2007) 293(1):H238–245. doi: 10.1152/ajpheart.01122.2006
- Lewis RS. Calcium signaling mechanisms in T lymphocytes. *Annu Rev Immunol* (2001) 19:497–521. doi: 10.1146/annurev.immunol.19.1.497
- Hogan PG. Calcium-NFAT transcriptional signalling in T cell activation and T cell exhaustion. *Cell Calcium* (2017) 63:66–9. doi: 10.1016/j.cecc.2017.01.014
- Beeton C, Wulff H, Barbaria J, Clot-Faybesse O, Pennington M, Bernard D, et al. Selective blockade of T lymphocyte k(+) channels ameliorates experimental autoimmune encephalomyelitis, a model for multiple sclerosis. *Proc Natl Acad Sci USA* (2001) 98(24):13942–7. doi: 10.1073/pnas.241497298
- Beeton C, Wulff H, Standifer NE, Azam P, Mullen KM, Pennington MW, et al. Kv1.3 channels are a therapeutic target for T cell-mediated autoimmune diseases. *Proc Natl Acad Sci USA* (2006) 103(46):17414–9. doi: 10.1073/pnas.0605136103
- Rus H, Pardo CA, Hu L, Darrah E, Cudrici C, Niculescu T, et al. The voltage-gated potassium channel Kv1.3 is highly expressed on inflammatory infiltrates in multiple sclerosis brain. *Proc Natl Acad Sci USA* (2005) 102(31):11094–9. doi: 10.1073/pnas.0501770102
- She G, Hou MC, Zhang Y, Zhang Y, Wang Y, Wang HF, et al. Gal-3 (Galectin-3) and KCa3.1 mediate heterogeneous cell coupling and myocardial fibrogenesis driven by β AR (β -adrenoceptor) activation. *Hypertension* (2020) 75(2):393–404. doi: 10.1161/HYPERTENSIONAHA.119.13696
- de Castro TBR, Canesso MCC, Boroni M, Chame DF, Souza D de L, de Toledo NE, et al. Differential modulation of mouse heart gene expression by infection with two trypanosoma cruzi strains: A transcriptome analysis. *Front Genet* (2020) 11:1031. doi: 10.3389/fgen.2020.01031
- Köberle F. The causation and importance of nervous lesions in American trypanosomiasis. *Bull World Health Organ* (1970) 42(5):739–43. doi: 10.1055/s-2008-1041127
- Villanueva MS. Trypanosomiasis of the central nervous system. *Semin Neurol* (1993) 13(2):209–18. doi: 10.1055/s-2008-1041127
- Amorim DS, Olsen EG. Assessment of heart neurons in dilated (congestive) cardiomyopathy. *Br Heart J* (1982) 47(1):11–8. doi: 10.1136/hrt.47.1.11
- de Souza MM, Andrade SG, Barbosa AA, Macedo Santos RT, Alves VA, Andrade ZA. Trypanosoma cruzi strains and autonomic nervous system pathology

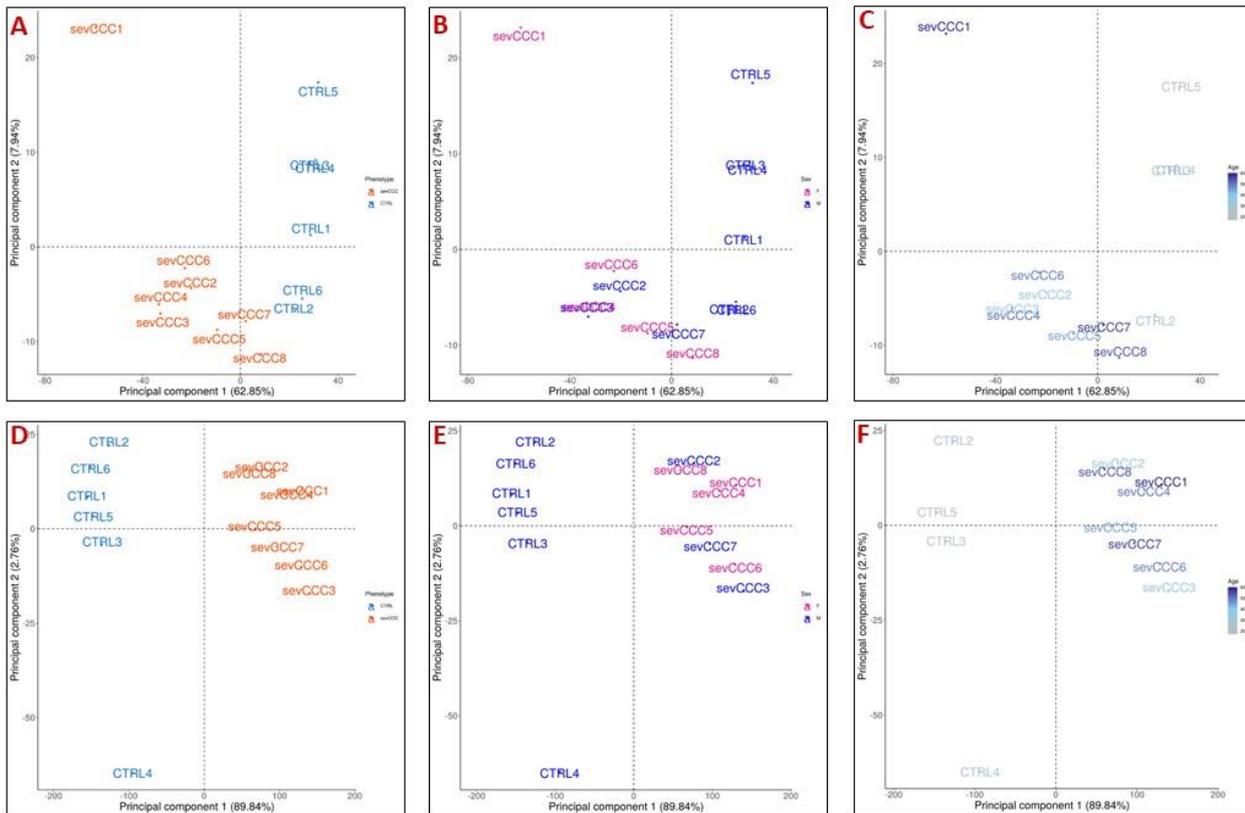
- in experimental chagas disease. *Mem Inst Oswaldo Cruz* (1996) 91(2):217–24. doi: 10.1590/S0074-02761996000200018
42. Machado CR, Caliarí MV, de Lana M, Tafuri WL. Heart autonomic innervation during the acute phase of experimental American trypanosomiasis in the dog. *Am J Trop Med Hygiene* (1998) 59(3):492–6. doi: 10.4269/ajtmh.1998.59.492
43. Machado CRS, Ribeiro ALP, Machado CRS, Ribeiro ALP. Experimental american trypanosomiasis in rats: sympathetic denervation, parasitism and inflammatory process. *Memórias do Instituto Oswaldo Cruz* (1989) 84(4):549–56. doi: 10.1590/S0074-02761989000400013
44. Machado MPR, Rocha AM, de Oliveira LF, de Cuba MB, de Loss IO, Castellano LR, et al. Autonomic nervous system modulation affects the inflammatory immune response in mice with acute chagas disease. *Exp Physiol* (2012) 97(11):1186–202. doi: 10.1113/expphysiol.2012.066431
45. Elenkov IJ, Wilder RL, Chrousos GP, Vizi ES. The sympathetic nerve—an integrative interface between two supersystems: the brain and the immune system. *Pharmacol Rev* (2000) 52(4):595–638. doi: 10.1189/jlb.1105625
46. Kin NW, Sanders VM. It takes nerve to tell T and B cells what to do. *J Leukoc Biol* (2006) 79(6):1093–104. doi: 10.1189/jlb.1105625
47. Kohm AP, Sanders VM. Norepinephrine and beta 2-adrenergic receptor stimulation regulate CD4+ T and B lymphocyte function *in vitro* and *in vivo*. *Pharmacol Rev* (2001) 53(4):487–525. doi: 10.1038/nri1810
48. Sternberg EM. Neural regulation of innate immunity: a coordinated nonspecific host response to pathogens. *Nat Rev Immunol* (2006) 6(4):318–28. doi: 10.1038/nri1810
49. Borda ES, Sterin-Borda L. Antiadrenergic and muscarinic receptor antibodies in chagas' cardiomyopathy. *Int J Cardiol* (1996) 54(2):149–56. doi: 10.1016/j.ijcard.2014.12.023
50. Bonney KM, Engman DM. Autoimmune pathogenesis of chagas heart disease. *Am J Pathol* (2015) 185(6):1537–47. doi: 10.1016/j.ajpath.2014.12.023
51. Machado FS, Dutra WO, Esper L, Gollob K, Teixeira MM, Factor SM, et al. Current understanding of immunity to trypanosoma cruzi infection and pathogenesis of chagas disease. *Semin Immunopathol* (2012) 34(6):753–70. doi: 10.1007/s00281-012-0351-7
52. Chevillard C, Nunes JPS, Frade AF, Almeida RR, Pandey RP, Nascimento MS, et al. Disease tolerance and pathogen resistance genes may underlie trypanosoma cruzi persistence and differential progression to chagas disease cardiomyopathy. *Front Immunol* (2018) 9:2791. doi: 10.3389/fimmu.2018.02791
53. O'Connell RM, Taganov KD, Boldin MP, Cheng G, Baltimore D. MicroRNA-155 is induced during the macrophage inflammatory response. *Proc Natl Acad Sci U S A* (2007) 104(5):1604–9. doi: 10.1073/pnas.0610731104
54. Corsten MF, Papageorgiou A, Verhesen W, Carai P, Lindow M, Obad S, et al. MicroRNA profiling identifies microRNA-155 as an adverse mediator of cardiac injury and dysfunction during acute viral myocarditis. *Circ Res* (2012) 111(4):415–25. doi: 10.1161/CIRCRESAHA.112.267443
55. Kanhere A, Hertweck A, Bhatia U, Gökmen MR, Perucha E, Jackson I, et al. T-Bet and GATA3 orchestrate Th1 and Th2 differentiation through lineage-specific targeting of distal regulatory elements. *Nat Commun* (2012) 3:1268. doi: 10.1038/ncomms2260
56. Kohu K, Ohmori H, Wong WF, Onda D, Wakoh T, Kon S, et al. The Runx3 transcription factor augments Th1 and down-modulates Th2 phenotypes by interacting with and attenuating GATA3. *J Immunol* (2009) 183(12):7817–24. doi: 10.4049/jimmunol.0802527
57. Talebi F, Ghorbani S, Chan WF, Boghazian R, Masoumi F, Ghasemi S, et al. MicroRNA-142 regulates inflammation and T cell differentiation in an animal model of multiple sclerosis. *J Neuroinflammation* (2017) 14:55. doi: 10.1186/s12974-017-0832-7
58. Nogueira LG, Santos RHB, Fiorelli AI, Mairena EC, Benvenuti LA, Bocchi EA, et al. Myocardial gene expression of T-bet, GATA-3, ror- γ t, FoxP3, and hallmark cytokines in chronic chagas disease cardiomyopathy: an essentially unopposed TH1-type response. *Mediators Inflamm* (2014) 2014:914326. doi: 10.1155/2014/914326
59. Chae WJ, Bothwell ALM. Canonical and non-canonical wnt signaling in immune cells. *Trends Immunol* (2018) 39(10):830–47. doi: 10.1016/j.it.2018.08.006
60. Volpini X, Ambrosio LF, Brajin MA, Brugo MB, Aoki MP, Rivarola HW, et al. Wnt signaling plays a key role in the regulation of the immune response and cardiac damage during trypanosoma cruzi infection. *ACS Infect Dis* (2021) 7(3):566–78. doi: 10.1021/acscinfed.0c00590
61. Volpini X, Ambrosio LF, Fozzatti L, Insfran C, Stempin CC, Cervi L, et al. Trypanosoma cruzi exploits wnt signaling pathway to promote its intracellular replication in macrophages. *Front Immunol* (2018) 9:859. doi: 10.3389/fimmu.2018.00859
62. *Wnt signaling pathways in myocardial infarction and the therapeutic effects of wnt pathway inhibitors | acta pharmacologica sinica*. Available at: <https://www.nature.com/articles/s41401-018-0060-4>.
63. Aisagbonhi O, Rai M, Ryzhov S, Atria N, Feoktistov I, Hatzopoulos AK. Experimental myocardial infarction triggers canonical wnt signaling and endothelial-to-mesenchymal transition. *Dis Model Mech* (2011) 4(4):469–83. doi: 10.1242/dmm.006510
64. Piersma B, Bank RA, Boersema M. Signaling in fibrosis: TGF- β , WNT, and YAP/TAZ converge. *Front Med (Lausanne)* (2015) 2:59. doi: 10.3389/fmed.2015.00059
65. Vallée A, Lecarpentier Y, Guillevin R, Vallée JN. Interactions between TGF- β 1, canonical WNT/ β -catenin pathway and PPAR γ in radiation-induced fibrosis. *Oncotarget* (2017) 8(52):90579–604. doi: 10.1038/ncomms1734
66. Akhmetshina A, Palumbo K, Dees C, Bergmann C, Venalis P, Zerr P, et al. Activation of canonical wnt signalling is required for TGF- β -mediated fibrosis. *Nat Commun* (2012) 3:735. doi: 10.1038/ncomms1734
67. Ferreira RR, de Souza EM, de Oliveira FL, Ferrão PM, Gomes LHF, Mendonça-Lima L, et al. Proteins involved on TGF- β pathway are up-regulated during the acute phase of experimental chagas disease. *Immunobiology* (2016) 221(5):587–94. doi: 10.1016/j.imbio.2016.01.009
68. Araújo-Jorge TC, Waghbi MC, Bailly S, Feige JJ. The TGF- β pathway as an emerging target for chagas disease therapy. *Clin Pharmacol Ther* (2012) 92(5):613–21. doi: 10.1038/clpt.2012.102
69. Li Y, Hiroi Y, Liao JK. Notch signaling as an important mediator of cardiac repair and regeneration after myocardial infarction. *Trends Cardiovasc Med* (2010) 20(7):228–31. doi: 10.1016/j.tcm.2011.11.006
70. Huang F, Zhu X, Hu XQ, Fang ZF, Tang L, Lu XL, et al. Mesenchymal stem cells modified with miR-126 release angiogenic factors and activate notch ligand delta-like-4, enhancing ischemic angiogenesis and cell survival. *Int J Mol Med* (2013) 31(2):484–92. doi: 10.3892/ijmm.2012.1200
71. *MicroRNA-384-5p promotes endothelial progenitor cell proliferation and angiogenesis in cerebral ischemic stroke through the delta-like ligand 4-mediated notch signaling pathway*. Available at:
72. Wu F, Yu B, Zhang X, Zhang Y. Cardioprotective effect of notch signaling on the development of myocardial infarction complicated by diabetes mellitus. *Exp Ther Med* (2017) 14(4):3447–54. doi: 10.3892/etm.2017.4932
73. Zhou XL, Fang YH, Wan L, Xu QR, Huang H, Zhu RR, et al. Notch signaling inhibits cardiac fibroblast to myofibroblast transformation by antagonizing TGF- β 1/Smad3 signaling. *J Cell Physiol* (2019) 234(6):8834–45. doi: 10.1002/jcp.27543
74. Yu B, Song B. Notch 1 signalling inhibits cardiomyocyte apoptosis in ischaemic postconditioning. *Heart Lung Circ* (2014) 23(2):152–8. doi: 10.1016/j.hlc.2013.07.004
75. Boopathy AV, Pendergrass KD, Che PL, Yoon YS, Davis ME. Oxidative stress-induced Notch1 signaling promotes cardiogenic gene expression in mesenchymal stem cells. *Stem Cell Res Ther* (2013) 4(2):43. doi: 10.1186/scrt190
76. Tu L, Fang TC, Artis D, Shestova O, Pross SE, Maillard I, et al. Notch signaling is an important regulator of type 2 immunity. *J Exp Med* (2005) 202(8):1037–42. doi: 10.1084/jem.20050923
77. Guzmán-Rivera D, Liempi A, González-Herrera F, Fuentes-Retamal S, Carrillo I, Abarca P, et al. Simvastatin improves cardiac function through notch 1 activation in BALB/c mice with chronic chagas cardiomyopathy. *Antimicrob Agents Chemother* (2020) 64(8):e02141–19. doi: 10.1371/journal.pntd.0009874
78. Casares-Marfil D, Kerick M, Andrés-León E, Bosch-Nicolau P, Molina IChagas Genetics CYTED Network, et al. GWAS loci associated with chagas cardiomyopathy influences DNA methylation levels. *PLoS Negl Trop Dis* (2021) 15(10):e0009874. doi: 10.1371/journal.pntd.0009874

COPYRIGHT

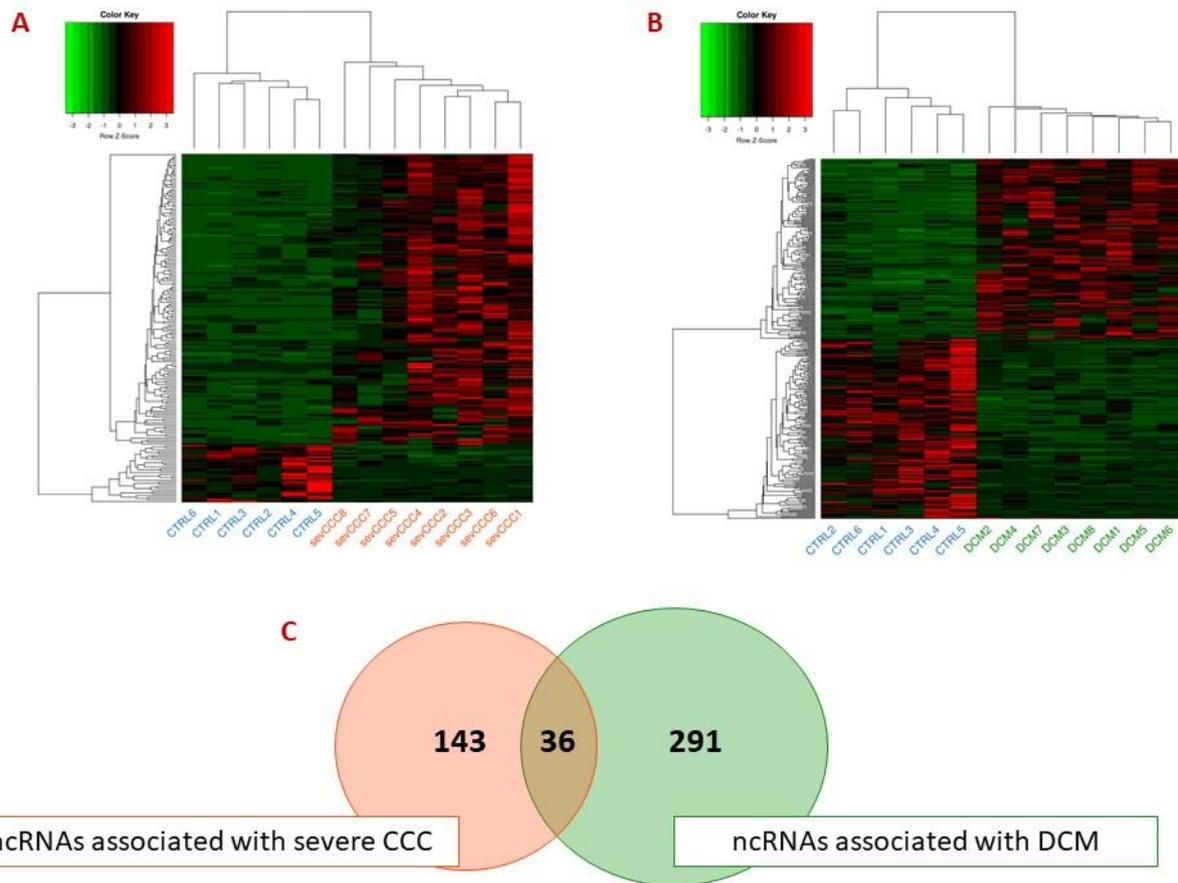
© 2022 Brochet, Ianni, Laugier, Frade, Silva Nunes, Teixeira, Mady, Ferreira, Ferré, Santos, Kuramoto, Cabantous, Steffen, Stolf, Pomerantz, Fiorelli, Bocchi, Pissetti, Saba, Cândido, Dias, Sampaio, Gaiotto, Marin-Neto, Fragata, Zaniratto, Siqueira, Peixoto, Rigaud, Bacal, Buck, Almeida, Lin-Wang, Schmidt, Martinelli, Hirata, Donadi, Costa Pereira, Rodrigues Junior, Puthier, Kalil, Spinelli, Cunha-Neto and Chevillard. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



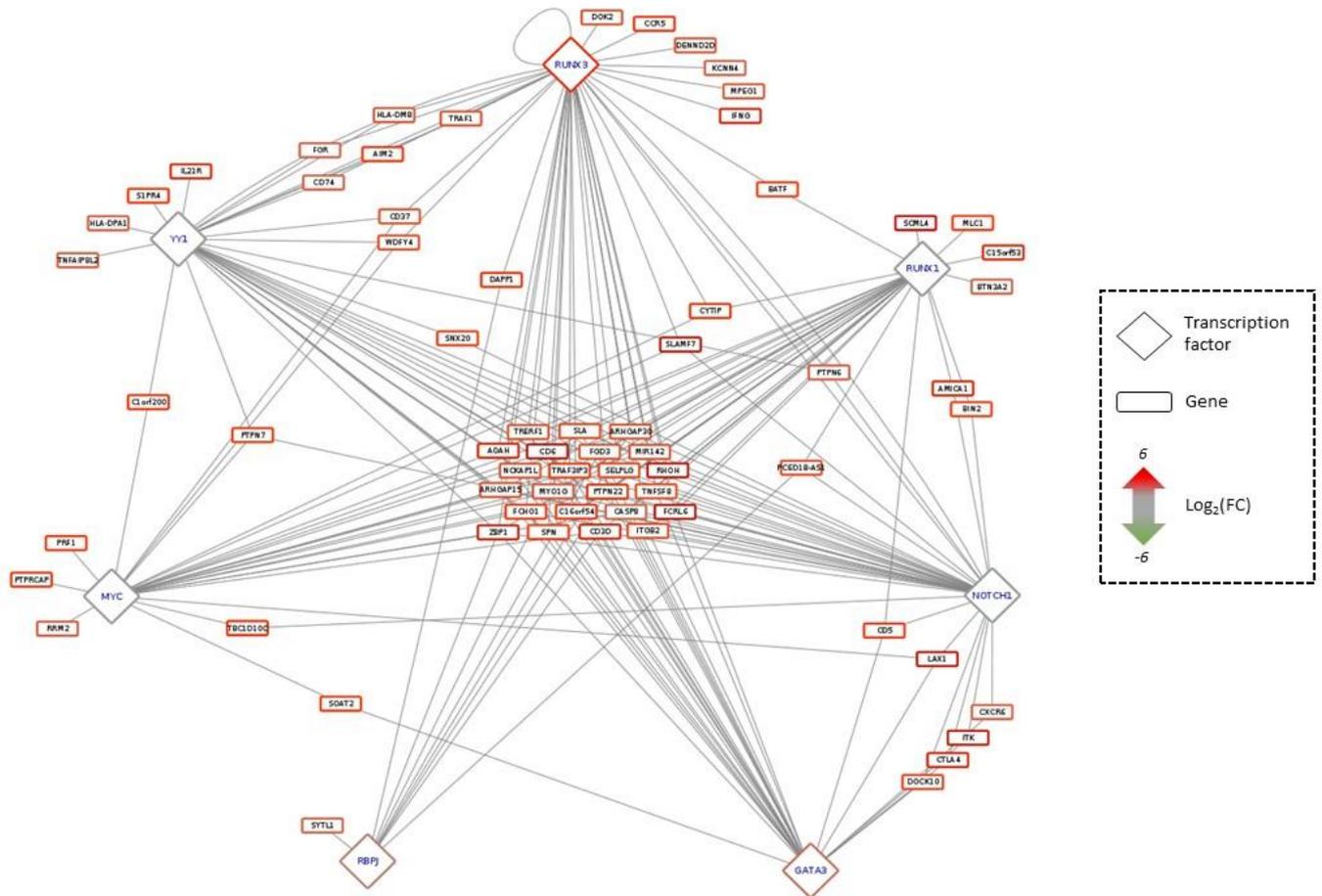
Supplementary Figure 1. Volcanopots of differential expression testing for CCC and DCM. Volcanoplot showing the difference of gene expression between **(A)** controls and CCC or **(B)** controls and DCM. On the x-axis is represented the $\log_2(\text{FC})$ and on the y-axis the $-\log_{10}(\text{adjusted pvalue})$. Overexpressed DEGs are in red, and underexpressed in green.



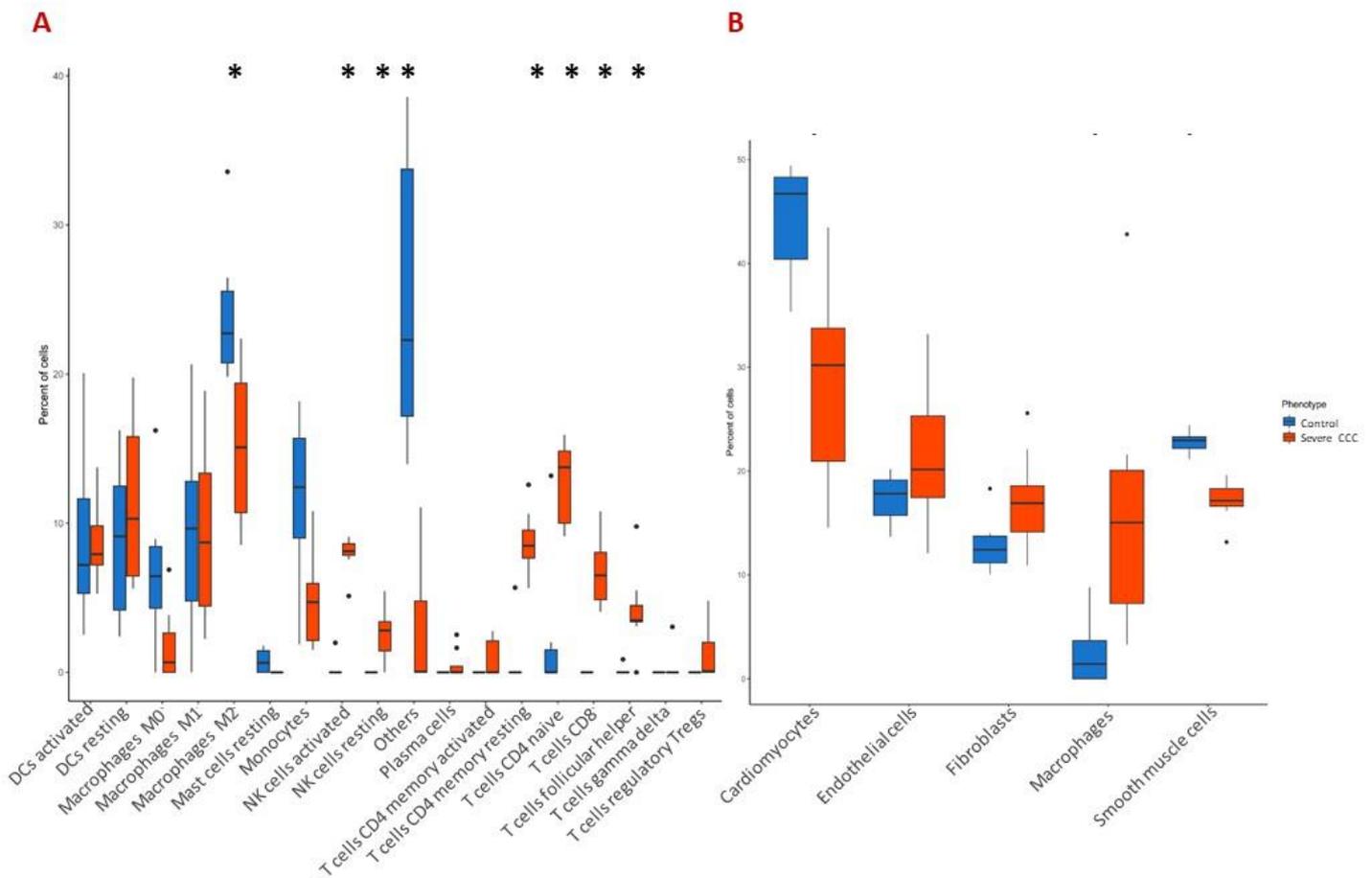
Supplementary Figure 2. Analysis of samples distribution in the space of differentially expressed genes or differentially methylated CpG sites between control and severe CCC. Age and sex impact on gene expression or on methylation. **(A)** Scatterplot of the two first principal component of a PCA of 6 Control and 8 Case samples executed in the space of the 1409 differentially expressed genes between the two conditions. **(B)** Scatterplot of the two first principal component of a PCA of 6 Control and 8 Case samples executed in the space of the 1409 differentially expressed genes between the two conditions colored by sex (blue: male, pink: female). **(C)** Scatterplot of the two first principal component of a PCA of 6 Control and 8 Case samples executed in the space of the 1409 differentially expressed genes between the two conditions colored by age (from grey to blue). **(D)** Scatterplot of the two first principal component of a PCA of 6 Control and 8 Case samples executed in the space of the 16883 differentially methylated CpG sites between the two conditions. **(E)** Scatterplot of the two first principal component of a PCA of 6 Control and 8 Case samples executed in the space of the 16883 differentially methylated CpG sites between the two conditions colored by sex (blue: male, pink: female). **(F)** Scatterplot of the two first principal component of a PCA of 6 Control and 8 Case samples executed in the space of the 16883 differentially methylated CpG sites between the two conditions colored by age (from grey to blue).



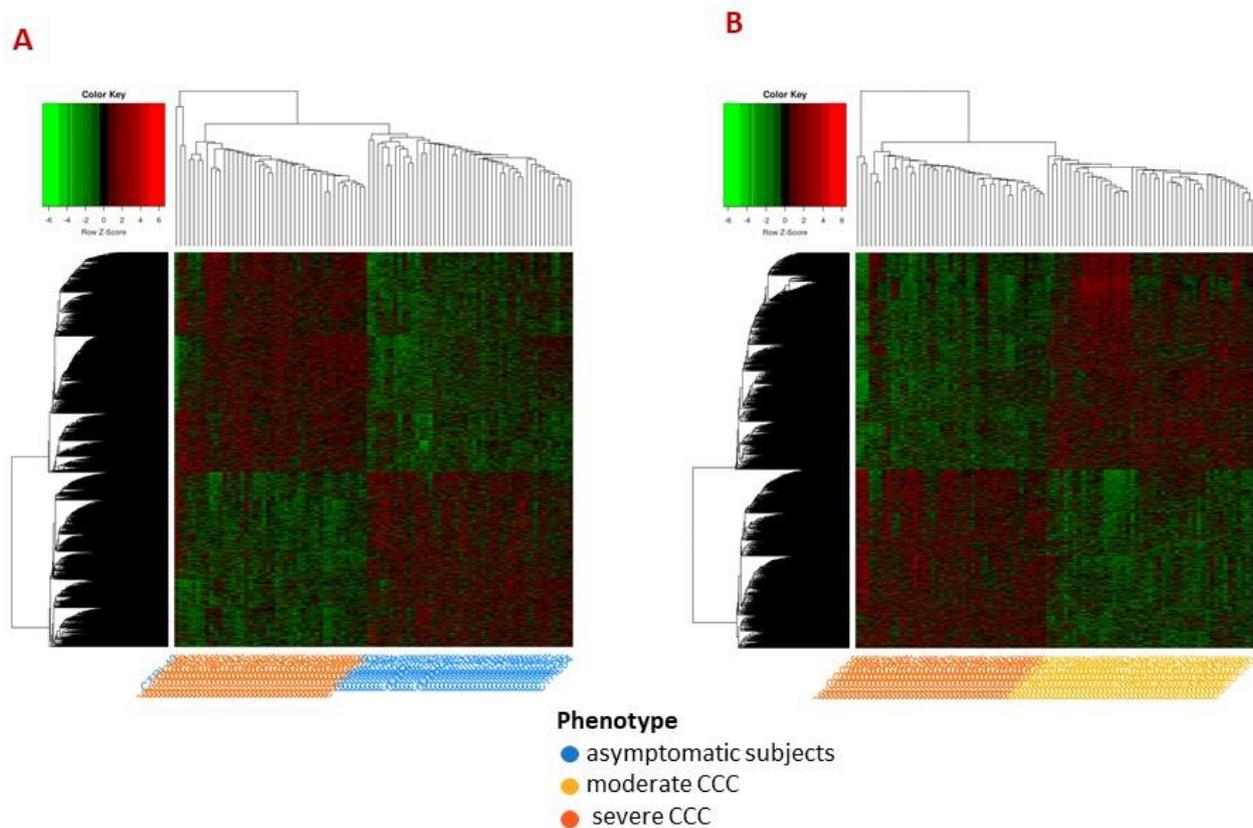
Supplementary Figure 3. ncRNA analysis in severe CCC and DCM. Control samples identifiers are written in blue, severe CCC samples identifiers in red and DCM samples in green. **(A)** Hierarchical Clustering Analysis (HCA) performed on 6 control and 8 severe CCC samples, based on expression of 179 differentially expressed ncRNAs. **(B)** Hierarchical Clustering Analysis (HCA) performed on 6 control and 8 DCM samples, based on expression of 327 differentially expressed ncRNAs. **(C)** Venn diagram of shared differentially expressed ncRNAs between severe CCC and DCM.



Supplementary Figure 4. Predicted regulatory interaction in cardiac muscular or nervous system process. Network composed by 6 transcription factors (TF) involved in cardiac muscular and/or nervous system process, and their 68 targeted genes, predicted by OLOGRAM according to ReMap database. TF are written in blue in diamond, and genes in black in rectangle. Shapes borders are colored according to the fold change, from green to red.



Supplementary Figure 5. Estimation of cell proportion in control and severe CCC heart tissue. Deconvolution of the RNA-seq bulk gene expressions of 6 control and 8 severe CCC samples to infer the proportion of cells in the samples. The * represent cell types whose proportion is significantly different (Wilcoxon test, FDR <0.05) between controls and cases. **(A)** Deconvolution using cell type signature with 22 immunological cell types (LM22 signature matrix). **(B)** Deconvolution using cell type signature with 5 left ventricle related cell types.



Supplementary Figure 6. Analysis of samples clustering based on differentially methylated CpG sites in blood samples. Asymptomatic samples identifiers are written in blue, moderate CCC in orange and severe CCC in red. **(A)** Hierarchical Clustering Analysis (HCA) performed on 48 asymptomatic and 90 severe CCC samples, based on expression of 12624 differentially methylated position. **(B)** Hierarchical Clustering Analysis (HCA) performed on 47 moderate CCC and 43 severe CCC samples, based on methylation level of 6735 differentially methylated position.

Toutes les tables supplémentaires sont directement **accessibles en ligne** à l'adresse suivante :

<https://www.frontiersin.org/articles/10.3389/fimmu.2022.958200/full>

2.2.2. Confirmation in silico de l'impact de la méthylation de l'ADN dans le développement des cardiomyopathies chagasiennes chroniques

Article 4 : Blood DNA methylation marks discriminate Chagas cardiomyopathy disease clinical forms.

Pauline Brochet, Barbara Ianni, João P S Nunes, Amanda F Frade, Priscila C Teixeira, Charles Mady, Ludmila R P Ferreira, Andreia Kuramoto, Cristina W Pissetti, Bruno Saba, Darlan D S Cândido, Fabrício Dias, Marcelo Sampaio, José A Marin-Neto, Abílio Fragata, Ricardo C F Zaniratto, Sergio Siqueira, Giselle D L Peixoto, Vagner O C Rigaud, Paula Buck, Rafael R Almeida, Hui Tzu Lin-Wang, André Schmidt, Martino Martinelli, Mario H Hirata, Eduardo Donadi, Virmondes Rodrigues Junior, Alexandre C Pereira, Jorge Kalil, Lionel Spinelli*, Edecio Cunha-Neto*, Christophe Chevillard*.

Front Immunol. 2022 Sep 29;13:1020572. doi: 10.3389/fimmu.2022.1020572. PMID: 36248819; PMCID: PMC9558220.

Présentation de la publication :

Cette étude fait directement suite à l'étude précédente, utilisant les données de méthylation obtenues à partir du sang de patients asymptomatiques et cardiaques. Ici, la sélection de variants n'a pas été faite à partir d'approches statistiques, mais en s'appuyant sur des modèles d'apprentissage automatique. La cohorte initiale a été divisée en deux jeux de données : les données d'entraînement (70% des échantillons) et les données de test (30% des échantillons). Tout d'abord, les CpG présentant au minimum 10% de variation entre les groupes ont été sélectionnés. Ensuite, différents modèles d'apprentissage automatique ont été employés sur la cohorte d'entraînement : arbre de décision, forêt aléatoire (random forest), régression logistique et SVM linéaire (Support-Vector Machine). Pour chacune de ces méthodes, les CpG ayant la plus grande contribution ont été conservés en utilisant la technique RFE (Recursive Caractéristique Elimination). L'entraînement du modèle a été réalisé sur la cohorte d'entraînement, et la prédiction sur celle de test, afin de confirmer les résultats obtenus sur un groupe indépendant.

Deux études ont été conduites en utilisant les données de méthylation, ayant pour objectif de prédire le phénotype des individus. Le but de ces analyses était, dans un premier temps, de prédire l'apparition des formes cardiaques à partir du sang de patients en phase chronique, et dans un deuxième temps, de prédire l'évolution des CCC en comparant les patients atteints de CCC modérée à ceux atteints de CCC sévère. Dans le premier cas, une précision de 95 % dans la prédiction du phénotype CCC a été atteinte en se basant sur la méthylation de 35 sites CpG, dont 20 se trouvent dans divers gènes associés à la réponse immunitaire, au système nerveux, au transport ionique ou au processus de l'ATP. En ce qui concerne l'évolution de la maladie, seulement 33 sites CpG étaient nécessaires pour prédire avec une précision de 96 % le phénotype modéré ou sévère. La plupart de ces sites sont situés dans des régions intergéniques, suggérant un rôle régulateur potentiel. Cependant, 15 sites CpG se localisent dans des gènes associés à la neurogenèse, aux filaments d'actine ou au transport ionique, similaires aux résultats obtenus précédemment.

Cette courte étude a permis d'identifier des sites de méthylation permettant de discriminer différentes formes chroniques de la maladie de Chagas. Ces informations pourraient conduire à la découverte de biomarqueurs pertinents. De plus, la capacité de la méthylation à prédire efficacement le phénotype de patients tend à confirmer son rôle important dans le développement de la pathologie.

Contribution personnelle :

J'ai réalisé l'intégralité des analyses présentes dans cette étude, dont je suis la première auteur. J'ai également rédigé intégralement la première version de l'article, et ai participé aux corrections.



OPEN ACCESS

EDITED BY
Ulrike Kemmerling,
University of Chile, Chile

REVIEWED BY
Juan Diego Maya,
University of Chile, Chile
Julio A. Urbina,
Instituto Venezolano de
Investigaciones Cientificas,
Venezuela
Israel Molina,
Vall d'Hebron University
Hospital, Spain

*CORRESPONDENCE
Christophe Chevillard
christophe.chevillard@univ-amu.fr
Edecio Cunha-Neto
edecunha@gmail.com
Lionel Spinelli
lionel.spinelli@univ-amu.fr

[†]These authors have contributed
equally to this work

SPECIALTY SECTION
This article was submitted to
Inflammation,
a section of the journal
Frontiers in Immunology

RECEIVED 16 August 2022
ACCEPTED 15 September 2022
PUBLISHED 29 September 2022

CITATION
Brochet P, Ianni B, Nunes JPS,
Frade AF, Teixeira PC, Mady C,
Ferreira LRP, Kuramoto A, Pissetti CW,
Saba B, Cândido DDS, Dias F,
Sampaio M, Marin-Neto JA, Fragata A,
Zaniratto RCF, Siqueira S, Peixoto GDL,
Rigaud VOC, Buck P, Almeida RR,
Lin-Wang HT, Schmidt A, Martinelli M,
Hirata MH, Donadi E, Rodrigues
Junior V, Pereira AC, Kalil J, Spinelli L,
Cunha-Neto E and Chevillard C (2022)
Blood DNA methylation marks
discriminate Chagas cardiomyopathy
disease clinical forms.
Front. Immunol. 13:1020572.
doi: 10.3389/fimmu.2022.1020572

Blood DNA methylation marks discriminate Chagas cardiomyopathy disease clinical forms

Pauline Brochet¹, Barbara Ianni², João P. S. Nunes^{1,2,3,4},
Amanda F. Frade^{2,3,4}, Priscila C. Teixeira^{2,3,4}, Charles Mady⁵,
Ludmila R. P. Ferreira⁶, Andreia Kuramoto², Cristina W. Pissetti⁷,
Bruno Saba⁸, Darlan D. S. Cândido^{2,3,4}, Fabrício Dias⁹,
Marcelo Sampaio⁸, José A. Marin-Neto⁹, Abílio Fragata⁸,
Ricardo C. F. Zaniratto², Sergio Siqueira¹⁰, Giselle D. L. Peixoto¹⁰,
Vagner O. C. Rigaud², Paula Buck¹¹, Rafael R. Almeida^{2,3,4},
Hui Tzu Lin-Wang⁸, André Schmidt⁹, Martino Martinelli¹⁰,
Mario H. Hirata¹², Eduardo Donadi⁹, Virmondos Rodrigues Junior⁷,
Alexandre C. Pereira¹¹, Jorge Kalil^{2,3,4}, Lionel Spinelli^{1*},
Edecio Cunha-Neto^{2,3,4*†} and Christophe Chevillard^{1*†}

¹Aix Marseille Univ, TAGC Theories and Approaches of Genomic Complexity, Institut MarMaRa, Marseille, France, ²Laboratory of Immunology, Heart Institute Instituto do Coração (InCor), School of Medicine, University of São Paulo, São Paulo, Brazil, ³Division of Clinical Immunology and Allergy, School of Medicine, University of São Paulo, São Paulo, Brazil, ⁴Instituto Nacional de Ciência e Tecnologia, INCT, iii- Institute for Investigation in Immunology, São Paulo, Brazil, ⁵Myocardopathies and Aortic Diseases Unit, Heart Institute, Instituto do Coração (InCor), School of Medicine, University of São Paulo, São Paulo, Brazil, ⁶RNA Systems Biology Laboratory (RSBL), Departamento de Morfologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil, ⁷Laboratory of Immunology, Universidade Federal Do Triângulo Mineiro (UFTM), Uberaba, Brazil, ⁸Laboratório de Investigação Molecular em Cardiologia, Instituto de Cardiologia Dante Pazzanese (IDPC), São Paulo, Brazil, ⁹School of Medicine of Ribeirão Preto (FMRP), University of São Paulo, Ribeirão Preto, Brazil, ¹⁰Pacemaker Clinic, Heart Institute Instituto do Coração (InCor), School of Medicine, University of São Paulo, São Paulo, Brazil, ¹¹Heart Institute Instituto do Coração (InCor), School of Medicine, University of São Paulo, São Paulo, São Paulo, Brazil, ¹²Department of Clinical and Toxicological Analyses, Faculty of Pharmaceutical Sciences, University of São Paulo (USP), São Paulo, Brazil

Chagas disease is a parasitic disease from South America, affecting around 7 million people worldwide. Decades after the infection, 30% of people develop chronic forms, including Chronic Chagas Cardiomyopathy (CCC), for which no treatment exists. Two stages characterized this form: the moderate form, characterized by a heart ejection fraction (EF) ≥ 0.4 , and the severe form, associated to an EF < 0.4 . We propose two sets of DNA methylation biomarkers which can predict in blood CCC occurrence, and CCC stage. This analysis, based on machine learning algorithms, makes predictions with more than 95% accuracy in a test cohort. Beyond their predictive capacity, these CpGs are located near genes involved in the immune response, the nervous system, ion transport or ATP synthesis, pathways known to be deregulated in CCCs. Among these genes, some are also differentially expressed in heart tissues.

Interestingly, the CpGs of interest are tagged to genes mainly involved in nervous and ionic processes. Given the close link between methylation and gene expression, these lists of CpGs promise to be not only good biomarkers, but also good indicators of key elements in the development of this pathology.

KEYWORDS

chagas disease, cardiomyopathy, blood, biomarkers, methylation

Introduction

Chagas disease is an endemic disease from South America, caused by a parasite, *Trypanosoma cruzi*, and affecting around 7 million people. With migration flow, this disease can now be found in non-endemic country, notably in North America (1) ($n > 300,000$), Europe (2) ($n > 100,000$), Japan (3) ($n > 4,000$) or Australia (4) ($n > 1,000$). After the infection, patients present an acute stage which is mostly asymptomatic (ASY). Then comes the chronic forms, where 70% of them remains asymptomatic, with no end organ damage (the so-called indeterminate stage). However, 30% develop Chagas disease Cardiomyopathy (CCC) (5). CCC had been divided in two stages based on heart ejection fraction: moderate CCC ($EF \geq 0.4$) and severe CCC ($EF < 0.4$) (6–8). Some drugs are effective on *T. cruzi*, but does not cure the CCC, reducing the parasitemia, without having any effect on heart damage (9). The only way out for CCC patients is the placement of a pacemaker, or a heart transplant. The early diagnosis of Chagas disease is therefore essential.

During the acute stage, disease diagnosis is commonly made by microscopy, considering the limited sensitivity of the direct test (10). However, in chronic stage, parasitemia is very low, or even null. The Pan American Health Organization (PAHO) recommends using two serological tests (two techniques based on different antigens) in parallel and, in case of discordant results, to perform these tests again on a new sample (11). If the results remain unclear, a confirmation test should be achieved (12). For CCC especially, an ECG and/or an echocardiogram is made to confirm cardiac involvement (13). BNP and NT-proBNP, well-known markers be associated with cardiac dysfunction (14), have been associated to Chagas cardiomyopathy (15, 16), but are not specific to this pathology. Others markers, including miRNAs (17, 18), cytokines (19) or metalloproteinases (20) have been proposed as biomarkers for CCC, but no confirmation has been made in a test cohort at this time. The only diagnosis of CCC currently in place is a clinical diagnosis, which is difficult to access for the most remote populations.

A previous analysis (21) has highlighted differences of DNA methylation in blood of asymptomatic and CCC patients. Moreover, some differences have also been demonstrated between moderate and severe CCC. Blood DNA methylation

has already been proposed as biomarker for several diseases (22–24). Here, we used machine learning methods on both asymptomatic and CCC blood DNA methylation data to predict Chagas disease, as well as Chagas disease stage.

Methods

Ethical considerations

The protocol was approved by the institutional review boards of the University of São Paulo School of Medicine and INSERM (French National Institute of Health and Medical Research). Written informed consent was obtained from all patients. All experimental methods comply with the Helsinki Declaration.

Blood DNA collection and DNA methylation analysis

Blood samples (5 to 15 ml of blood) from CCC patients were collected in EDTA tubes. Genomic DNA was isolated using standard salted methods and the methylation analysis was done using the same protocol as tissue DNAs.

Blood DNA methylation data

138 patients were selected randomly from our Chagas bank. It included 48 asymptomatic subjects, 46 moderate CCC patients and 44 severe CCC patients (Supplementary Table 1). The age and sex ratio were not significantly different between the 3 groups (age mean and ratio female/male for all phenotypes: asymptomatic: age: 57.63, ratio =1; moderate: age: 56.89, ratio=1.14; severe: age: 59.59, ratio=0.95). In a second time, these 138 samples were randomly distributed between the training (70%) and validation (30%) cohorts. This random distribution was done in such a way that the age and sex ratio was still not so different between the groups in the two sub-cohorts (Training cohort (age mean and ratio female/male for all

phenotypes: asymptomatic: age: 62.45, ratio=1; moderate: age: 60.13, ratio=1.14; severe: age: 57.18, ratio=0.86), validation cohort (age and ratio female/male for all phenotypes: asymptomatic: age: 52.82, ratio=1; moderate: age: 53.67, ratio=1.13; severe: age: 62, ratio=1). The methylation data are available under the reference: (GEO accession: GSE191082).

Biomarker identification for disease forms

Since data contains a lot of features (736,661), feature selection was performed in two steps, on the training group only. The scripts used for the following steps are available on Github (<https://github.com/TAGC-ComplexDisease/biomarkersChagas>). First, the delta beta (difference of beta means) was computed between the two phenotypes of interest. Only the CpGs having at least 10% methylation differences were retained. Then, a machine learning (ML) analysis was done in Python with Scikit-learn library. Four supervised ML methods were considered: decision tree, random forest, logistic regression and linear SVM (Support-Vector Machine, a linear classifier). For each method, recursive feature elimination (RFE) was performed, and the best model (best accuracy) with the minimal set of feature was selected using 10-time cross-validation. Finally, model parameters were optimized with a grid search to obtain the final prediction on the validation group.

Results

Symptomatic cardiac form prediction

After feature selection based on delta beta values, 86 CpGs were selected. Among all the tested models, linear SVM seems to have the better prediction on training dataset with the minimal number of features (Figure 2A). According to this analysis, the

SVM was trained with 35 features (Supplementary Table 2). The model parameters optimization was performed using a grid search where the L2 penalty varies between 0.01 and 10. Finally, with a L2 penalty of 1, 42 of 44 patients phenotype of the validation dataset were correctly predicted (accuracy = 0.95), with a sensitivity of 0.96 and a specificity of 0.94 (area under the curve: 0.996) (Figure 1A). Those 35 features are mainly located in the body of genes (n=20), or in intergenic regions (n=11). Particularly, 3 CpGs are located in LHX6, and 3 in POU6F2. All those genes are involved in biological process associated to Chagas disease: nervous system (LHX6, POU6F2, MDGA1, DISC1, PCSK9), immune system (ZMIZ, HLA-DRB1), Wnt pathway (DISC1), ion transport (KCNK15, PCSK9), striated muscles (SMYD3) or ATP metabolic process (ATP5S).

Chagas cardiomyopathy stage prediction

After feature selection based on delta beta, 108 CpGs were selected. Among all the tested models, random forest seems to have the better prediction on training dataset with the minimal number of features (Figure 2B). According to this analysis, the SVM was trained with 33 features (Supplementary Table 3). The model parameters optimization was performed using a grid search where the number of estimators varies between 50 and 200. Finally, 150 estimators, 27 of 28 patients phenotype (accuracy = 0.96) of the validation dataset were correctly predicted, with a sensitivity of 1 and a specificity of 0.93 (area under the curve: 1) (Figure 1B). Those 33 features are mainly located in 18 intergenic regions. Other CpGs are located in 15 genes, and more precisely in 6 gene body and 7 promoter regions. Here, genes are involved in various biological processes, from ion transport (KCNK1, MFI2), actin filament (PACSIN1), generation of neurons (TNN, PACSIN1) or MAPK cascade (DUSP22). 2 CpGs are in common with those used as biomarker between ASY and CCC: cg24000535 (LOC101928909) and cg21873524 (intergenic).

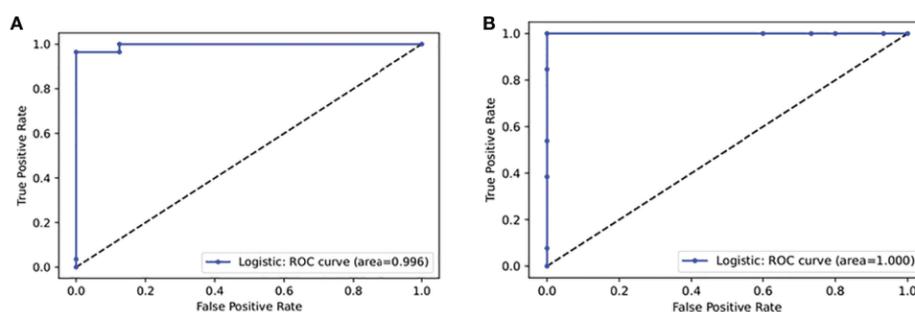
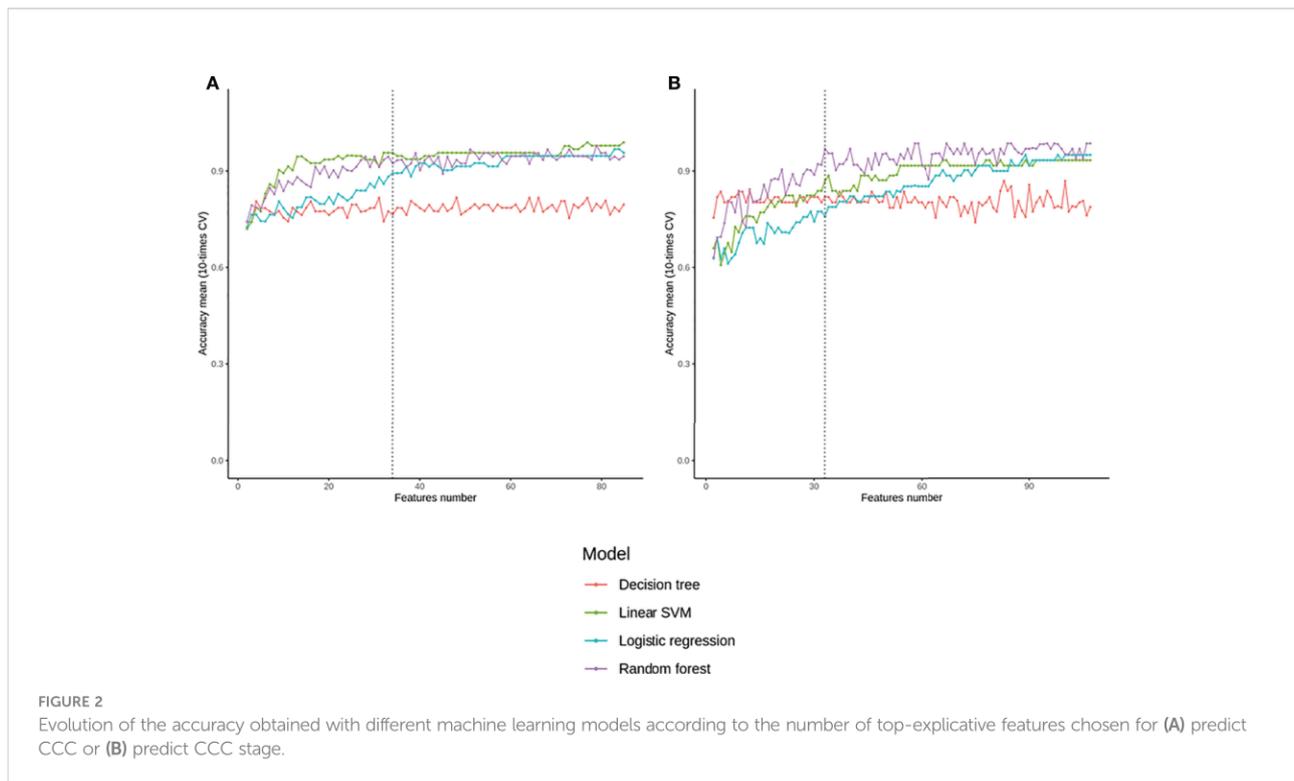


FIGURE 1
Receiver Operating Characteristic (ROC) curves produced by (A) linear SVM to predict CCC on 44 patients and (B) random forest to predict CCC stage on 28 patients.



Discussion

Alterations of heart tissue DNA methylation profiles have been associated to the development of dilated cardiomyopathies (25, 26) and chagas diseases (21, 27). Recently, we have studied the DNA methylation in the blood of asymptomatic, moderate and severe CCC by hypothesizing that the blood data reflect the phenotype. We had found 12624 DMPs (Differentially Methylated Position) between asymptomatic and severe CCC blood samples and 6735 CpGs were found as DMPs between moderate and severe CCC.

In our study, based on machine learning approaches, we have identified 35 CCC-specific methylation markers. Those CpGs could distinguish controls (asymptomatic) to CCC from blood samples with 96% of sensitivity and 94% of specificity in independent validation sets. 3 of them (cg02872767, cg24540763, cg25134647) are also differentially methylated between asymptomatic and severe CCC in heart tissue (21). Similarly, 33 CpGs have been identified, allowing to predict the progression of this pathology (from moderate to severe CCC), with a sensitivity of 100% and a specificity of 93%. In conclusion, we identified two set of methylation markers potentially useful for Chagas disease diagnostic. The first one permit to discriminate patients with Chagas cardiomyopathy from asymptomatic patients, with 95% of precision. The

second allows to predict Chagas cardiomyopathy severity stage, according to the heart ejection fraction rate, with 96% of precision.

Interestingly, most of these markers were not differentially in heart tissue of patients. The main message of this report is the finding that peripheral blood epigenetic marks are good markers of clinical form, implying that epigenetic events are closely related to CCC progression. These markers are highlighting the same biological processes that have been associated to the disease development such as ion transport, ATP metabolic process, immune system, Wnt system, nervous systems, striated muscles and actin filament. These findings are important as these are under-lighting biological pathways that will have to be targeted in drug design.

To propose large-scale reproducible biomarkers, a consensus Target Product Profile (TPP) has been developed for Chagas disease (28) stipulating that marker should be able to detect the effects of drug treatments, be detectable with limited resources and not vary according to the strain of the parasite. Given the high specificity of these assays, these methylation sites appear to be good candidates to decipher the pathogenic process, or to be used as blood biomarkers, and further studies will be necessary to potentially validate their possible use in the clinic, in accordance with the TPP consensus.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/[Supplementary Material](#).

Ethics statement

The studies involving human participants were reviewed and approved by INSERM IRB University of Sao Paulo. The patients/participants provided their written informed consent to participate in this study.

Author contributions

Study design: PBr, JK, LS, ECN, CC. Phenotype characterization: BI, CM, SS, CWP, BS, FD, MS, JAMN, AF, GDLP, PBu, HTL-W, AS, MM, MHH, ED, ACP, VRJ. Experimental analysis: PBr, AFF, JPSN, PCT, LRPF, AK, DDSC, RCFZ, VOGR, RRA. Statistical analysis: PBr, LS, ECN, CC. Manuscript preparation: PBr, LS, ECN, CC.

Funding

This work was supported by the Institut National de la Santé et de la Recherche Médicale (INSERM); the Aix-Marseille University (grant number: AMIDEX “International_2018” MITOMUTCHAGAS); the French Agency for Research (Agence Nationale de la Recherche-ANR (grant numbers: “Br-Fr-Chagas”, “landscardio”); the CNPq (Brazilian Council for Scientific and Technological Development); and the FAPESP (São Paulo State Research Funding Agency Brazil (grant numbers: 2013/50302-3, 2014/50890-5); the National Institutes of Health/USA (grant numbers: 2 P50 AI098461-02 and 2U19AI098461-06). This work was founded by the Inserm

References

- Bern C, Messenger LA, Whitman JD, Maguire JH. Chagas Disease in the United States: a Public Health Approach. *Clin Microbiol Rev* (2019) 33. doi: 10.1128/CMR.00023-19
- Antinori S, Galimberti L, Bianco R, Grande R, Galli M, Corbellino M. Chagas disease in Europe: A review for the internist in the globalized world. *Eur J Intern Med* (2017) 43:6–15. doi: 10.1016/j.ejim.2017.05.001
- Imai K, Maeda T, Sayama Y, Osa M, Mikita K, Kurane I, et al. Chronic Chagas disease with advanced cardiac complications in Japan: Case report and literature review. *Parasitol Int* (2015) 64:240–2. doi: 10.1016/j.parint.2015.02.005
- Jackson Y, Pinto A, Pett S. Chagas disease in Australia and New Zealand: risks and needs for public health interventions. *Trop Med Int Health* (2014) 19:212–8. doi: 10.1111/tmi.12235

Cross-Cutting Project GOLD. This project has received funding from the Excellence Initiative of Aix-Marseille University - A*Midex a French “Investissements d’Avenir programme”- Institute MarMaRa AMX-19-IET-007. JN was a recipient of a MarMaRa fellowship. EC-N, JK, ALR and ECS are recipients of productivity awards by CNPq. The funders did not play any role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgments

Center de Calcul Intensif d’Aix-Marseille is acknowledged for granting access to its high performance computing resources.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2022.1020572/full#supplementary-material>

- Pérez-Molina JA, Molina I. Chagas disease. *Lancet* (2018) 391:82–94. doi: 10.1016/s0140-6736(17)31612-4
- Frade AF, Teixeira PC, Ianni BM, Pissetti CW, Saba B, Wang LH, et al. Polymorphism in the alpha cardiac muscle actin 1 gene is associated to susceptibility to chronic inflammatory cardiomyopathy. *PLoS One* (2013) 8: e83446. doi: 10.1371/journal.pone.0083446
- Nogueira LG, Santos RH, Ianni BM, Fiorelli AI, Mairena EC, Benvenuti LA, et al. Myocardial chemokine expression and intensity of myocarditis in Chagas cardiomyopathy are controlled by polymorphisms in CXCL9 and CXCL10. *PLoS Negl Trop Dis* (2012) 6:e1867. doi: 10.1371/journal.pntd.0001867
- van Veldhuisen DJ, Linssen GC, Jaarsma T, van Gilst WH, Hoes AW, Tijssen JG, et al. B-type natriuretic peptide and prognosis in heart failure patients with

preserved and reduced ejection fraction. *J Am Coll Cardiol* (2013) 61:1498–506. doi: 10.1016/j.jacc.2012.12.044

9. Morillo CA, Marin-Neto JA, Avezum A, Sosa-Estani S, Rassi AJr., Rosas F, et al. Randomized Trial of Benznidazole for Chronic Chagas' Cardiomyopathy. *N Engl J Med* (2015) 373:1295–306. doi: 10.1056/NEJMoa1507574

10. Balouz V, Agüero F, Buscaglia CA. Chagas Disease Diagnostic Applications: Present Knowledge and Future Steps. *Adv Parasitol* (2017) 97:1–45. doi: 10.1016/bs.apar.2016.10.001

11. Organization PAH. *Guidelines for the diagnosis and treatment of Chagas disease*. Washington, D.C: Pan American Health Organization (PAHO) (2019).

12. Lapa JS, Saraiva RM, Hasslocher-Moreno AM, Georg I, Souza AS, Xavier SS, et al. Dealing with initial inconclusive serological results for chronic Chagas disease in clinical practice. *Eur J Clin Microbiol Infect Dis* (2012) 31:965–74. doi: 10.1007/s10096-011-1393-9

13. Benck L, Kransdorf E, Patel J. Diagnosis and Management of Chagas Cardiomyopathy in the United States. *Curr Cardiol Rep* (2018) 20:131. doi: 10.1007/s11886-018-1077-5

14. Chow SL, Maisel AS, Anand I, Bozkurt B, de Boer RA, Felker GM, et al. Role of Biomarkers for the Prevention, Assessment, and Management of Heart Failure: A Scientific Statement From the American Heart Association. *Circulation* (2017) 135:e1054–e91. doi: 10.1161/CIR.0000000000000490

15. Echeverría LE, Rojas LZ, Gomez-Ochoa SA, Rueda-Ochoa OL, Sosa-Vesga CD, Muka T, et al. Cardiovascular biomarkers as predictors of adverse outcomes in chronic Chagas cardiomyopathy. *PLoS One* (2021) 16:e0258622. doi: 10.1371/journal.pone.0258622

16. Brito BOF, Pinto-Filho MM, Cardoso CS, Di Lorenzo Oliveira C, Ferreira AM, de Oliveira LC, et al. Association between typical electrocardiographic abnormalities and NT-proBNP elevation in a large cohort of patients with Chagas disease from endemic area. *J Electrocardiol* (2018) 51:1039–43. doi: 10.1016/j.jelectrocard.2018.08.031

17. Gomez-Ochoa SA, Bautista-Nino PK, Rojas LZ, Hunziker L, Muka T, Echeverría LE. Circulating MicroRNAs and myocardial involvement severity in chronic Chagas cardiomyopathy. *Front Cell Infect Microbiol* (2022) 12:922189. doi: 10.3389/fcimb.2022.922189

18. Linhares-Lacerda L, Granato A, Gomes-Neto JF, Conde L, Freire-de-Lima L, de Freitas EO, et al. Circulating Plasma MicroRNA-208a as Potential Biomarker of Chronic Indeterminate Phase of Chagas Disease. *Front Microbiol* (2018) 9:269. doi: 10.3389/fmicb.2018.00269

19. De Alba-Alvarado M, Salazar-Schettino PM, Jimenez-Alvarez L, Cabrera-Bravo M, Garcia-Sancho C, Zenteno E, et al. Th-17 cytokines are associated with severity of *Trypanosoma cruzi* chronic infection in pediatric patients from endemic areas of Mexico. *Acta Trop* (2018) 178:134–41. doi: 10.1016/j.actatropica.2017.11.009

20. Medeiros NI, Gomes JAS, Fiuzza JA, Sousa GR, Almeida EF, Novaes RO, et al. MMP-2 and MMP-9 plasma levels are potential biomarkers for indeterminate

and cardiac clinical forms progression in chronic Chagas disease. *Sci Rep* (2019) 9:14170. doi: 10.1038/s41598-019-50791-z

21. Brochet P, Ianni B, Laugier L, Frade AF, Nunes JPS, Teixeira P, et al. Epigenetic regulation of transcription factor binding motifs promotes Th1 response in Chagas disease Cardiomyopathy. *Front Immunol* (2022). doi: 10.3389/fimmu.2022.958200

22. Fransquet PD, Lacaze P, Saffery R, Phung J, Parker E, Shah R, et al. Blood DNA methylation signatures to detect dementia prior to overt clinical symptoms. *Alzheimers Dement (Amst)* (2020) 12:e12056. doi: 10.1002/dad2.12056

23. Lange CP, Campan M, Hinoue T, Schmitz RF, van der Meulen-de Jong AE, Slingerland H, et al. Genome-scale discovery of DNA-methylation biomarkers for blood-based detection of colorectal cancer. *PLoS One* (2012) 7:e50266. doi: 10.1371/journal.pone.0050266

24. Shu C, Justice AC, Zhang X, Marconi VC, Hancock DB, Johnson EO, et al. DNA methylation biomarker selected by an ensemble machine learning approach predicts mortality risk in an HIV-positive veteran population. *Epigenetics* (2021) 16:741–53. doi: 10.1080/15592294.2020.1824097

25. Jo BS, Koh IU, Bae JB, Yu HY, Jeon ES, Lee HY, et al. Methyloome analysis reveals alterations in DNA methylation in the regulatory regions of left ventricle development genes in human dilated cardiomyopathy. *Genomics* (2016) 108:84–92. doi: 10.1016/j.ygeno.2016.07.001

26. Haas J, Frese KS, Park YJ, Keller A, Vogel B, Lindroth AM, et al. Alterations in cardiac DNA methylation in human dilated cardiomyopathy. *EMBO Mol Med* (2013) 5:413–29. doi: 10.1002/emmm.201201553

27. Laugier L, Frade AF, Ferreira FM, Baron MA, Teixeira PC, Cabantous S, et al. Whole-Genome Cardiac DNA Methylation Fingerprint and Gene Expression Analysis Provide New Insights in the Pathogenesis of Chronic Chagas Disease Cardiomyopathy. *Clin Infect Dis* (2017) 65:1103–11. doi: 10.1093/cid/cix506

28. Porras AI, Yadon ZE, Altchek J, Britto C, Chaves GC, Flevaud L, et al. Target Product Profile (TPP) for Chagas Disease Point-of-Care Diagnosis and Assessment of Response to Treatment. *PLoS Negl Trop Dis* (2015) 9:e0003697. doi: 10.1371/journal.pntd.0003697

COPYRIGHT

© 2022 Brochet, Ianni, Nunes, Frade, Teixeira, Mady, Ferreira, Kuramoto, Pissetti, Saba, Cândido, Dias, Sampaio, Marin-Neto, Fragata, Zaniratto, Siqueira, Peixoto, Rigaud, Buck, Almeida, Lin-Wang, Schmidt, Martinelli, Hirata, Donadi, Rodrigues Junior, Pereira, Kalil, Spinelli, Cunha-Neto and Chevillard. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Toutes les tables supplémentaires sont directement **accessibles en ligne** à l'adresse suivante :

<https://www.frontiersin.org/articles/10.3389/fimmu.2022.1020572/full>

3. Exploration multi-omiques des cardiomyopathies chagasiques chroniques : de la collecte de données à l'intégration approfondie

3.1. Centralisation des résultats omiques de la maladie de Chagas dans une seule base de données

Article 5 : ChagasDB: 80 years of publicly available data on the molecular host response to *Trypanosoma cruzi* infection in a single database.

Pauline Brochet, Jean-Christophe Mouren, Laurent Hannouche, Fabrice Lopez, Benoit Ballester, Edecio Cunha-Neto*, Lionel Spinelli*, Christophe Chevillard*.

Database (Oxford). 2023 May 23;2023:baad037. doi: 10.1093/database/baad037. PMID: 37221041; PMCID: PMC10205463.

Présentation de la publication :

Depuis la découverte de la maladie de Chagas, de nombreuses analyses omiques ont été mises en place, permettant de décrypter le processus pathogénique. Cette étude a eu pour objectif d'identifier tous les articles présentant ces résultats, et de les regrouper dans une base de données accessible à tous. En se concentrant uniquement sur les caractéristiques significativement associées à cette maladie, 193 articles ont été retenus. Ces derniers totalisent 19 phénotypes différents, dont 14 associés à l'infection par *T. cruzi* (phase aiguë, phase chronique, formes asymptomatiques, digestives, cardiaques, cardiaques modérées, cardiaques sévères...) et 5 comme contrôles (sains, cardiomyopathies dilatées...). De plus, 24 souches de *T. cruzi* différentes ont pu être utilisées dans ces analyses, et 5 espèces différentes (homme, souris, rat, brebis et chien). Les CCC représentent 73% des articles répertoriés, illustrant l'intérêt pour cette maladie.

Une grande partie des gènes et protéines associées à la maladie de Chagas a été retrouvée en commun dans tous les organismes, bien que l'homme et la souris présentent également un grand nombre de caractéristiques spécifiques.

Malgré tout, dans ces deux organismes, les mêmes processus biologiques sont affectés, incluant la réponse immunitaire, le métabolisme mitochondrial ou le système nerveux.

Cette base de données, appelée ChagasDB, est disponible à l'adresse suivante : <https://chagasdb.tagc.univ-amu.fr>. Le site web permet d'explorer directement ces informations ou de les télécharger pour une analyse plus approfondie. Il propose également aux auteurs de soumettre directement leurs résultats.

Contribution personnelle :

J'ai mis en place le projet ChagasDB, une idée que je nourrissais depuis longtemps et qui remonte à avant le commencement de ma thèse. J'ai effectué la sélection des études et la création de la base de données. De plus, j'ai participé à la mise en place du site web. J'ai également rédigé intégralement la première version de l'article, dont je suis la première auteur, et ai participé aux corrections.

ChagasDB: 80 years of publicly available data on the molecular host response to *Trypanosoma cruzi* infection in a single database

Pauline Brochet ¹, Jean-Christophe Mouren ¹, Laurent Hannouche ², Fabrice Lopez ¹,
Benoit Ballester ¹, Edécio Cunha-Neto ^{3,4,5,*†}, Lionel Spinelli ^{1,*†} and
Christophe Chevillard ^{1,*†}

¹Institut National de la Santé Et de la Recherche Médicale (INSERM), Unité Mixte de Recherche (UMR)_1090, Aix Marseille Université, TAGC Theories and Approaches of Genomic Complexity, Institut MarMaRa, Marseille 13288, France

²C2VN, INSERM, INRA, Aix Marseille Université, Marseille 13005, France

³Laboratory of Immunology, Heart Institute Instituto do Coração (InCor), University of São Paulo, School of Medicine, São Paulo 05403-900, Brazil

⁴Division of Clinical Immunology and Allergy, University of São Paulo, School of Medicine, São Paulo 05403-900, Brazil

⁵Instituto Nacional de Ciência e Tecnologia, INCT, III- Institute for Investigation in Immunology, São Paulo 05403-900, Brazil

*Corresponding author: Tel: +551126615914/5906; Fax: +551126615953; Email: edecunha@gmail.com

Correspondence may also be addressed to Lionel Spinelli. Tel: +33491828735; Fax: +33491828701; Email: lionel.spinelli@univ-amu.fr and Christophe Chevillard. Tel: +33491828736; Fax: +33491828701; Email: christophe.chevillard@univ-amu.fr

[†]These authors contributed equally to this work.

Citation details: Brochet, P., Mouren, J., Hannouche, L. *et al.* ChagasDB: 80 years of publicly available data on the molecular host response to *Trypanosoma cruzi* infection in a single database. *Database* (2023) Vol. 2023: article ID baad037; DOI: <https://doi.org/10.1093/database/baad037>

Abstract

Chagas disease is a parasitological disease caused by *Trypanosoma cruzi* which affects ~7 million people worldwide. Per year, ~10 000 people die from this pathology. Indeed, ~30% of humans develop severe chronic forms, including cardiac, digestive or neurological disorders, for which there is still no treatment. In order to facilitate research on Chagas disease, a manual curation of all papers corresponding to 'Chagas disease' referenced on PubMed has been performed. All deregulated molecules in hosts (all mammals, humans, mice or others) following *T. cruzi* infection were retrieved and included in a database, named ChagasDB. A website has been developed to make this database accessible to all. In this article, we detail the construction of this database, its contents and how to use it.

Database URL: <https://chagasdb.tagc.univ-amu.fr>

Introduction

Chagas disease is a parasitological disease caused by a protozoan, *Trypanosoma cruzi*. This parasite is endemic from South America and contaminates multiple mammal species through a vector of the *Reduviidae* family or indirect contamination (food contaminated, congenital transmission...) (1). After infection with *T. cruzi*, infected mammals first present an acute phase. This phase is mostly asymptomatic but sometimes leads to physical signs, such as Romana's sign, as well as fever or headache. During this phase, a strong immune response is set up, characterized by an overproduction of Th1 lymphocytes, which will produce interferon gamma and destroy the parasite, albeit not providing sterilizing immunity (2). After several weeks, the infected mammals enter the chronic phase, which is mostly asymptomatic. However, severe forms can develop, notably cardiac, digestive, cardiodigestive or nervous. In humans, ~30% of individuals in the chronic phase

will develop these complications, the most frequent being chronic Chagas cardiomyopathy. Chagas disease affects ~7 million of people, and per year, ~10 000 people die from this pathology (3).

To understand the development of Chagas disease and to find effective treatments, many research teams have worked not only on patient samples but also on animal models (4, 5), such as mice, rats, dogs, hamsters or even sheep. Over the years, these studies have revealed numerous deregulated molecules in the hosts, associated with the infection of the parasite. With the advancement of technology, increasing number of features has been characterized, like mutations, genes, proteins, hormones or other chemical molecules. However, there is currently no database containing all this information. It is therefore difficult to make the links between new results and what has already been identified by the community.

To facilitate the sharing of knowledge related to Chagas disease and thus research on this disease, we have reviewed

Received 24 January 2023; Revised 22 April 2023; Accepted 28 April 2023

© The Author(s) 2023. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

193 articles from 1945 to 2022 and collected information on molecules in hosts associated with *T. cruzi* infection. The collected information has been summarized in a freely accessible online database, ChagasDB (<https://chagasdb.tagc.univ-amu.fr>), and as a downloadable file. Thus, any molecule of interest can easily be found, regardless of the organism, its type or the strain of *T. cruzi* responsible for its dysregulation. These data can then be used to enrich other studies. In this article, we detail the construction criteria of this database, as well as its content and how to use it.

Materials and methods

Article review

We selected the article matching the search ‘Chagas disease’ in PubMed. The selection of the articles was made following several steps based on different criteria. In the first time, the title of each article was read, and all articles not related to research on host responses to *T. cruzi* infection were discarded. Therefore, any review, epidemiological analysis or study of the parasite or insect vector has not been included in ChagasDB. The second step of article selection was performed based on the abstract. Only unmodified hosts (not knocked out or stimulated with a particular chemical compound) were retained. This can be *in vivo* studies or *in vitro* studies with unmodified cells. Finally, the third and last steps were applied, reading the ‘Material and Methods’ section, to verify that the article meets the established criteria. When selected, the complete article was parsed to review the results. Analyses showing significant associations between *T. cruzi* infection and host molecules were included in the database.

For each molecule of interest, the following information was retrieved: its name as used in the article, its type, its level of variation, the compared phenotypes, the analyzed tissue, the method of analysis, the organism, the strain of *T. cruzi* (if known), the Digital Object Identifier (DOI) and PubMed reference number (PMID) of the article, its title and its year of publication.

Database annotation

The collected features were annotated using different reference databases, depending on the species. For all genes-related features (protein, methylation site, mutation...), the name, description and ID of the associated genes were reported. In that case, for all of them, Ensembl (6) database was used. For *Homo sapiens* specifically, HUGO Gene Nomenclature Committee (7) and GeneCards (8) helped to identify some alias. microRNADaBase (miRBase) (9) and LNCipedia (10) were used for noncoding RNA. The Mouse Genome Informatics (11) and Rat Genome Database (12) were employed for *Mus musculus* and *Rattus norvegicus*, respectively.

Website development

The platform has been developed using a Python (v3.9.13) environment with the Django (v4.0.7) framework coupled to a SQLite (v3.39.3) database. The web interface makes use of JQuery (v3.6.0), HTML5 and the framework TailwindCSS (v3.2.4). ChagasDB was deployed with Docker (v20.10.12) and Docker Compose (v1.25.0). A CLI (command-line interface) tool using Typer (v0.4.1) was developed to ensure an easy and fast update of the platform. All the scripts used to develop the website are available at <https://github.com/TAGC-ComplexDisease/ChagasDB>.

ChagasDB features analysis

To get an overview of the information contained in the ChagasDB, a functional enrichment analysis was applied using R language. Due to the low number of features identified in *R. norvegicus*, *Canis lupus familiaris* and *Ovis aries*, only human and mouse features were considered for Gene Ontology (GO) biological process enrichment. Deregulated genes and proteins were analyzed using gProfileR (v0.7.0) (13), and top GOs were recovered using rrvigo (v1.8.0) (14). Finally, the results were illustrated using ggplot2. To compare genes or proteins between different species, orthologous genes were found using biomaRt R package (v2.52.0) (15).

Results and discussion

Among the 21 121 articles listed in PubMed as ‘Chagas disease’, we have retained 193 articles (Figure 1) that correspond to our selection criteria: experimental or computational analysis that found significant deregulation at molecular levels on host infected by *T. cruzi*. Although the oldest paper referenced in PubMed dates from 1945, the 193 articles in ChagasDB were published between 1995 and 2022, due to the lack of statistical test before 1995 (Supplementary Table S1). Information for each feature in the database was gathered, comprising a set of 25 distinct categories. These categories included details related to the feature itself, such as its name, description, ID and type, as well as information regarding the study conducted, including tissue, phenotypes, analysis, organism, population and *T. cruzi* strain. Additionally, the paper’s DOI, PMID, title and publication year were also included in the information set. All the details are available on Supplementary Table S2.

For now, 19 phenotypes are available in ChagasDB: five as *T. cruzi*-uninfected control (with several phenotypes) and 14 as cases, with different forms of Chagas disease. The controls comprise healthy controls and non-chagasic cardiomyopathy, including dilated, idiopathic and ischemic cardiomyopathy. In *T. cruzi*-infected phenotype, all disease stages are available, including acute stage, asymptomatic chronic stage and symptomatic chronic stage: cardiac (moderate, mild, severe or all), digestive, cardiodigestive or mixed. Some congenital infection and heart transplant chagasic patients were also included. Concerning the parasite, 24 *T. cruzi* strains were included in the database, corresponding to three discrete typing units, TcI, TcII and TcVI. For TcI were used Colombia, G, Sylvio, SylvioX10/4, AC, CA-I, Col1.8G2, K98 and Type 1 strains. TcII include ABC, Berenice-78, Y, JG, Lucky and SGO-Z12 strains. Finally, TcVI comprise CL-14, CL-Brener, Cvd, Q501/3, Tulahuén, Tehuantepec, RA and VD strains.

From the 193 articles in ChagasDB, 174 present experimental analysis of molecules associated with *T. cruzi* infection, while 33 present computational analysis. Most of them were interested in *H. sapiens* response to *T. cruzi* (142), but some animal models were also employed: *M. musculus* (40), *R. norvegicus* (12), *O. aries* (1) and *C. lupus familiaris* (1). As a result, a greater number and diversity of molecules have been identified in humans compared to other mammals (Table 1).

However, a large percent of genes and proteins deregulated are shared between mammals (Figure 2), with 55% of mouse, 69% of rat and 27% of human features shared between at least two species.

Beyond the deregulated genes shared between humans and mice, similar biological functions are affected at different

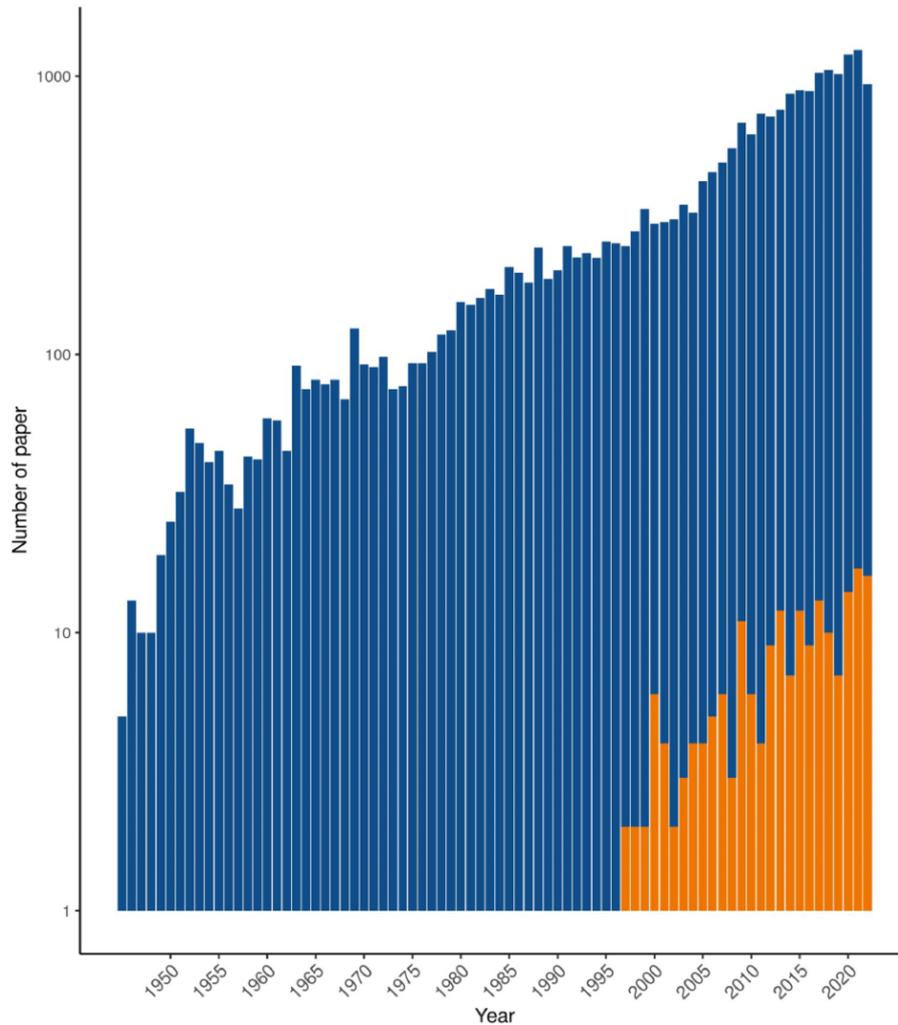


Figure 1. Number of paper references as 'Chagas disease' in PubMed (blue) and included in ChagasDB (orange) per year.

Table 1. Number of features included in ChagasDB by species

| | <i>Homo sapiens</i> | <i>Mus musculus</i> | <i>Rattus norvegicus</i> | <i>Canis lupus familiaris</i> | <i>Ovis aries</i> |
|--|---------------------|---------------------|--------------------------|-------------------------------|-------------------|
| RNA | 10 534 | 5885 | 289 | 11 | 3 |
| Alternative splicing | 1778 | | | | |
| Methylation | 41 041 | | | | |
| Transcription factor binding site (TFBS) | 30 | | | | |
| Polymorphism | 49 | | | | |
| Haplogroup | 6 | | | | |
| Protein | 339 | 56 | 15 | 3 | |
| Protein activity | 7 | 7 | | | |
| Chemical molecule | 49 | 2 | 5 | | |
| Electrolyte | 1 | 2 | | | |
| Hormone | 8 | | | | |

stages of the disease (Figure 3). In both species, the immune response and muscle-related process are affected in all disease stages, whereas nitric oxide process seems to be specific to acute stage. Interestingly, in *H. sapiens*, the immune response is not deregulated in asymptomatic stage, unlike acute and chronic symptomatic stages. Due to the low number of features associated with other phenotypes (digestive, cardioidigestive...), these phenotypes were not taken into account in the analysis (Supplementary Table S3).

While acute and asymptomatic forms of studies represent, respectively, 51 (26%) and 77 (40%) articles, cardiac forms totalize 141 articles (73%) of the whole 193 articles, illustrating the interest for this specific disease presentation. There is an urgent need to understand why it develops in order to find a treatment. The development of databases makes it possible to gather a large amount of information that is directly and easily accessible, thus facilitating research. Moreover, the pooling of results, which are sometimes difficult to

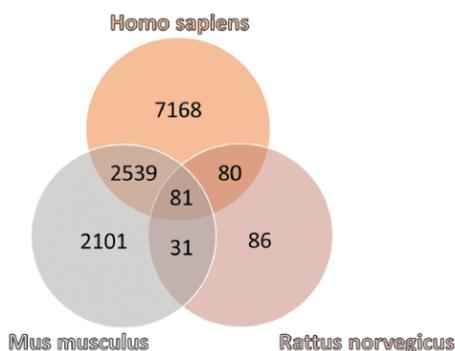


Figure 2. Venn diagram of the number of genes and proteins referenced in ChagasDB in *Homo sapiens*, *Mus musculus* and *Rattus norvegicus*.

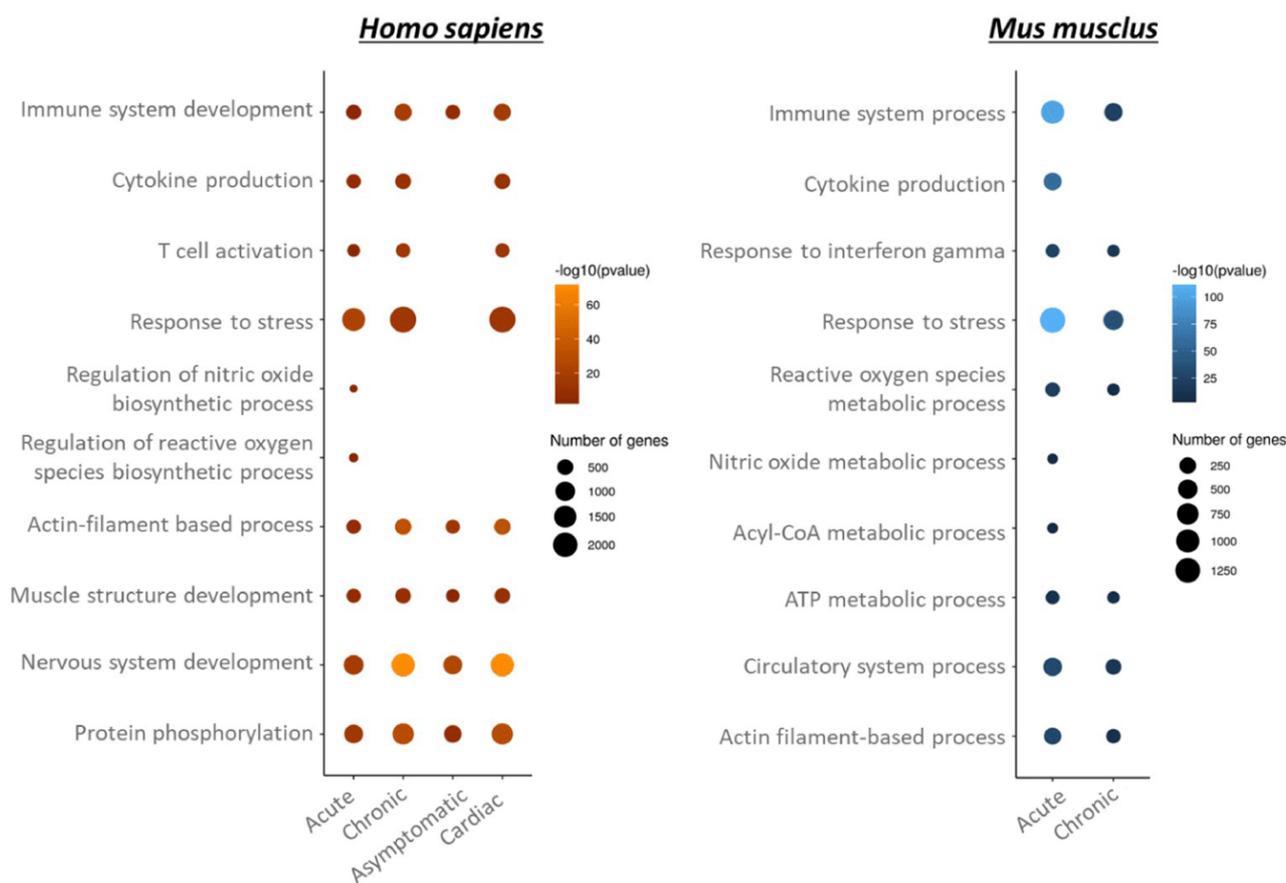


Figure 3. GO biological process enrichment of genes and proteins deregulated in *Homo sapiens* and *Mus musculus* per disease stage.

access, makes it possible to at least confirm one's own results and at best save time and resources. For this purpose, we have developed a website allowing not only to explore ChagasDB through queries but also to download the whole database. Briefly, users can enter a feature name, ID or description in the 'Search' bar tool, and the different corresponding features will appear. They can download the information relative to the features of interest. Users can apply some filters on different fields (like organism, *T. cruzi* strain...) to have access to a restricted number of results that correspond to their needs. Finally, users can also submit their own research results to improve the ChagasDB content.

To conclude, ChagasDB is a database gathering all the molecules identified in the literature as deregulated in hosts

due to *T. cruzi* infection. These data are easily accessible via a website, to be integrated in different analyses, or to be used as a confirmation. To our knowledge, this is the first database related to molecular response to Chagas disease. Currently, ChagasDB is limited to host molecules, but in the future, it could be extended to other types of data: cellular deregulations, responses to stimuli or treatments. Its objective will always be to facilitate research on Chagas disease, in order to help achieve a complete cure.

Supplementary material

Supplementary material is available at *Database* online.

Data availability

The ChagasDB can be downloaded by everybody without preliminary authorization.

Author contributions

P.B. and L.H. contributed to database; P.B. contributed to data curation; J.-C.M. and F.L. contributed to website; P.B. contributed to writing—original draft preparation; B.B., E.C.-N., L.S. and C.C. contributed to writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding

The Institut National de la Santé et de la Recherche Médicale (INSERM); the Aix Marseille University (AMIDEX ‘International_2018’ MITOMUTCHAGAS); the French Agency for Research (Agence Nationale de la Recherche-ANR (‘Br-Fr-Chagas’, ‘landscardio’); the INSERM Cross-Cutting Project GOLD; the Excellence Initiative of Aix Marseille University - A*Midex a French ‘Investissements d’Avenir programme’ - Institute MarMaRa AMX-19-IET-007; FAPESP (São Paulo State Research Funding Agency Brazil) (2022/00758-0); Brazilian National Research Council (Productivity Award to E.C.-N); the CAPES-COFECUB program (Me987/23).

Acknowledgement

French Ministry of Higher Education and Research PhD Fellowship (to J.-C.M. and P.B.). Center de Calcul Intensif d’Aix Marseille is acknowledged for hosting ChagasDB.

Feedback

The ChagasDB team thanks our future users for their future feedback to make ChagasDB useful for the community. Please contact for development requests (pauline.brochet@gmail.com or christophe.chevillard@univ-amu.fr).

Conflict of interest statement

The authors declare no conflict of interest.

References

1. Pérez-Molina, J.A. and Molina, I. (2018) Chagas disease. *Lancet*, **391**, 82–94.
2. Cristovão-Silva, A.C., Brelaz-de-castro, M.C.A., Hernandez, M.Z. et al. (2021) Chagas disease: Immunology of the disease at a glance. *Cytokine Growth Factor Rev.*, **62**, 15–22.
3. Lee, B.Y., Bacon, K.M., Bottazzi, M.E. et al. (2013) Global economic burden of Chagas disease: a computational simulation model. *Lancet Infect Dis*, **13**, 342–348.
4. Scharfstein, J., Gomes, J.D.A.S. and Correa-Oliveira, R. (2009) Back to the future in Chagas disease: from animal models to patient cohort studies, progress in immunopathogenesis research. *Mem. Inst. Oswaldo Cruz*, **104 Suppl 1**, 187–198.
5. Chatelain, E. and Scandale, I. (2020) Animal models of Chagas disease and their translational value to drug development. *Expert Opin. Drug Discov.*, **15**, 1381–1402.
6. Cunningham, F., Allen, J.E., Allen, J. et al. (2022) Ensembl 2022. *Nucleic Acids Res.*, **50**, D988–D995.
7. Tweedie, S., Braschi, B., Gray, K. et al. (2020) Genenames.org: the HGNC and VGNC resources in 2021. *Nucleic Acids Res.*, **49**, D939–D946.
8. Stelzer, G., Rosen, R., Plaschkes, I. et al. (2016) The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. *Current Protocols in Bioinformatics*, **54**, 1–30.
9. Kozomara, A., Birgaoanu, M. and Griffiths-Jones, S. (2019) miR-Base: from microRNA sequences to function. *Nucleic Acids Res.*, **47**, D155–D162.
10. Volders, P.-J., Anckaert, J., Verheggen, K. et al. (2019) LNCipedia 5: towards a reference set of human long non-coding RNAs. *Nucleic Acids Res.*, **47**, D135–D139.
11. Blake, J.A., Baldarelli, R., Kadin, J.A. et al. (2020) Mouse Genome Database (MGD): knowledgebase for mouse–human comparative biology. *Nucleic Acids Res.*, **49**, D981–D987.
12. Smith, J.R., Hayman, G.T., Wang, S.-J. et al. (2020) The Year of the Rat: the Rat Genome Database at 20: a multi-species knowledgebase and analysis platform. *Nucleic Acids Res.*, **48**, D731–D742.
13. Reimand, J., Arak, T., Adler, P. et al. (2016) g:Profiler—a web server for functional interpretation of gene lists (2016 update). *Nucleic Acids Res.*, **44**, W83–W89.
14. Sayols, S. (2023) rrvgo: a Bioconductor package for interpreting lists of Gene Ontology terms. *MicroPubl. Biol.*, **2023**, 10.17912/micropub.biology.000811.
15. Durinck, S., Spellman, P.T., Birney, E. et al. (2009) Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protoc.*, **4**, 1184–1191.

Toutes les tables supplémentaires sont directement **accessibles en ligne** à l'adresse suivante :

<https://academic.oup.com/database/article/doi/10.1093/database/baad037/7176384>

3.2. L'intégration multi-omiques a mis en évidence des mécanismes pathologiques uniques chez des patients atteints de cardiomyopathie chagasiqne chronique

Article 6 : Exploring global and specific pathogenic mechanisms in Chronic Chagas Cardiomyopathy through multi-omics integration

Pauline Brochet, Jorge Kalil, Vincent Procaccio, Edecio Cunha-Neto*, Lionel Spinelli*, Christophe Chevillard*.

En préparation, disponible sur medRxiv.

Présentation de la publication :

Les nombreuses études omiques précédemment conduites sur la maladie de Chagas puis menées au cours de ce projet ont mis en évidence la contribution de nombreux types de molécules dans le processus pathogénique associé aux CCC. Nous avons également démontré l'impact de différents mécanismes épigénétiques conduisant potentiellement à une dérégulation génétique. Par ailleurs, nous avons mis en évidence une contribution de la génétique mitochondriale au développement de cette maladie. A partir des données obtenues sur du tissu cardiaque de patients atteints de CCC sévères et de contrôles sains, nous avons mené une intégration multi-omiques, incluant transcriptomique, méthylomique, miRnomique et séquençage de la mitochondrie.

En raison du nombre limité d'échantillons et de données manquantes dans certaines omiques, nous avons d'abord appliqué plusieurs étapes de prétraitement des données, notamment l'imputation et la génération de données synthétiques. Ensuite, nous avons utilisé l'outil MOFA+ pour effectuer l'intégration proprement dite. Des comparaisons menées avec d'autres méthodes d'augmentation de données et d'intégration multi-omiques ont confirmé la pertinence de nos résultats. De plus, l'imputation des échantillons manquants n'a pas eu d'incidence significative sur les résultats obtenus après l'intégration, car la majorité d'entre eux étaient systématiquement retrouvés en supprimant aléatoirement un échantillon réel et en procédant à sa réimputation.

Afin d'interpréter les caractéristiques sélectionnées par MOFA+, deux réseaux de corrélation, spécifiques à chaque phénotype (contrôles sains et CCC) ont été créés. L'étude de ces réseaux a montré des mécanismes régulatoires spécifiques aux CCC, notamment trois clusters associés à différentes fonctions biologiques (cluster 2 : Lymphocytes T/TCR, clusters 1 et 3 : BCR et TRL). Ces résultats nous ont conduit à effectuer une analyse du répertoire des lymphocytes B et T, qui a révélé une forte convergence dans les chaînes légères du BCR, synonymes d'une pression de sélection continue.

Après avoir filtré les réseaux de corrélation à l'aide de bases de données publiques pour limiter les interactions peu probables, une analyse détaillée des clusters a été effectuée. Les gènes du cluster 2 sont principalement associés à la réponse inflammatoire, régulée principalement par le facteur de transcription IRF4, tandis que les clusters 1 et 3 montrent une implication partielle dans la régulation de la réponse immunitaire, principalement médiée par le TLR7. De nombreux ARN non codants ont été identifiés, impliqués notamment dans l'inflammation et les maladies cardiaques. De plus, nous avons repéré des sites de méthylation susceptibles d'influencer la fixation de facteurs de transcription spécifiques, pouvant ainsi contribuer à la dérégulation génétique constatée.

Lors d'une seconde intégration multi-omiques, effectuée exclusivement sur les patients atteints de CCC, des processus biologiques distincts ont été révélés chez 5 d'entre eux. Étant donné le nombre limité de gènes uniques à la plupart des patients, nous avons concentré notre attention sur l'étude des caractéristiques de deux échantillons : CCC3 et CCC6. CCC3 se caractérise par une forte expression de gènes liés à l'inflammation, notamment l'IFN- γ , ainsi qu'une expression élevée du facteur de transcription IRF4. En revanche, CCC6 présente une altération de l'expression de gènes associés aux processus mitochondriaux, en particulier au complexe I de la chaîne respiratoire mitochondriale, et est marqué par la présence de plusieurs mutations. La comparaison des profils d'expression des caractéristiques spécifiques à ces deux patients a mis en lumière des mécanismes distincts observés dans les réseaux précédemment établis.

En conclusion, notre étude a permis de partiellement reconstituer une partie du processus pathogénique associé aux CCC. L'analyse approfondie menée exclusivement sur les patients CCC a révélé des processus distincts, chacun étant observable dans le réseau construit précédemment. Ainsi, plusieurs mécanismes pathogéniques peuvent être impliqués, contribuant au développement du même phénotype. Il est important de noter que l'étude des réseaux est toujours en cours, et que de nouveaux résultats pourront émerger ultérieurement.

Contribution personnelle :

Ma contribution dans cet article, dont je suis la première auteur, a été majoritaire. J'ai pris en charge l'intégralité des analyses relatives à l'intégration multi-omiques, des données brutes aux réseaux de corrélation. J'ai également rédigé intégralement la première version de l'article.

Exploring global and specific pathogenic mechanisms in Chronic Chagas Cardiomyopathy through multi-omics integration

Pauline Brochet¹, Jorge Kalil^{2,3,4}, Vincent Procaccio⁵, Edecio Cunha-Neto^{2,3,4,*c}, Lionel Spinelli^{1,6,*c}, Christophe Chevillard^{1,*c}

Abstract

Chagas disease is a neglected disease from South America caused by a parasite, *Trypanosoma cruzi*. While most of infected people remains asymptomatic, around 30% develop Chronic Chagas Cardiomyopathy (CCC), a very lethal cardiomyopathy characterized by an exacerbated inflammatory response. The last few years, our team has set up multiple omics analysis. Briefly, we have pointed the over-expression of many genes involved in the Th1 lymphocyte response, as well as some epigenetic features potentially involved in their regulation, including miRNA, lncRNA and methylation site. Moreover, some mitochondria mutation seems to predispose to the development of CCC. In order to understand and characterize the impact of genetic and epigenetic elements on the pathogenic process associated to CCC, we have performed here a multi-omics integration, combining transcriptomic, methylomic, miRNomic and mitochondria sequencing. We have identified two distinct pathogenic pathways that vary among patients with chronic Chagas cardiomyopathy (CCC). One pathway is primarily influenced by IRF4, a transcription factor known for its involvement in the development of both B and T cells, while the other is driven by TLR signaling. Notably, genes related to B cells play a role in both of these processes. Additionally, we have detected certain similarities in the B cell receptors of all CCC patients, which may potentially contribute to autoimmunity. While further analysis is necessary to validate these findings, they collectively enhance our understanding of the pathogenic mechanisms associated with CCC.

¹Institut National de la Santé Et de la Recherche Médicale (INSERM), Unité Mixte de Recherche (UMR)_1090, Aix Marseille Université, TAGC Theories and Approaches of Genomic Complexity, Institut MarMaRa, Marseille, France

²Laboratory of Immunology, Heart Institute Instituto do Coração (InCor), University of São Paulo, School of Medicine, São Paulo, Brazil

³Division of Clinical Immunology and Allergy, University of São Paulo, School of Medicine, São Paulo, Brazil

⁴Instituto Nacional de Ciência e Tecnologia, INCT, III- Institute for Investigation in Immunology, São Paulo, Brazil

⁵Department of Genetics, University Hospital of Angers, 49933 Angers, France

⁶Aix Marseille Université, CNRS, INSERM, Centre d'Immunologie de Marseille-Luminy, Marseille, France

*Authors have an equal contribution.

^cCorresponding authors.

Introduction

Chagas disease is a parasitical disease caused by *Trypanosoma cruzi*, a parasite endemic from South America, and spreading across the globe with the migratory flow (1–3). After the infection, the acute phase is mostly asymptomatic, and most of the contaminated peoples remains asymptomatic (ASY) through the years. Nonetheless, approximately 30% develop Chronic Chagas Cardiomyopathy (CCC), the main cause of death of Chagas disease, which kill 10.000 peoples each year. At this stage, the parasitemia is very low, and anti-parasite treatments have no impact on heart deterioration (4).

CCC is characterized by a myocardial dilatation and an infiltrate of immune cells (mainly lymphocytes and macrophages), inducing continuous inflammation (5). Although the mechanisms driving the development of this cardiomyopathy remains unclear, there is a notable disparity in the differentiation of Th lymphocytes into Th1 and Th2 cells in CCC, which results in a prevalent Th1 profile in CCC patients and a high production of pro-inflammatory cytokines, including IFN- γ and TNF- α . Conversely, asymptomatic (ASY) patients present a more equilibrium lymphocytes Th profiles, with a higher production of anti-inflammatory cytokines like IL-10 (6). Furthermore, mitochondrial dysfunctions have been detected in CCC, leading to reduced mitochondrial activity and the consequent production of Reactive Oxygen Species (ROS). While during the acute phase, ROS serve to eliminate parasites, their prolonged accumulation in cardiac tissue eventually leads to tissue damage (7). The hypothesis of an autoimmune reaction developing in CCC patients has been proposed, supported by the identification of auto-reactive peptides that stimulate B or T cells in human sera (8).

With advances in large-scale analysis techniques, many study have focused on understanding the pathogenic process leading to the development of CCC (9). To this end, our team has carried out genomic, transcriptomic and methylomic studies highlighting processes specific to CCC patients. Focusing on mitochondrial DNA, only a limited number of mutations linked to cardiac diseases have been identified. However, there is evidence that certain mitochondrial haplogroups are

overrepresented among CCC patients. Transcriptomic studies have revealed the involvement of immune-related genes, including Toll-like receptors, Th-1 lymphocyte genes, and cytokine signaling, when compared to healthy asymptomatic, dilated cardiomyopathy (DCM) controls (10,11). Interestingly, some genes associated with cardiac and nervous system processes also exhibit dysregulation between CCC and DCM, illustrating the specificity of this disease. Some associations have been established between the expression of these genes and epigenetic regulatory elements, including non-coding genes such as miRNA (12) and lncRNA (13), and methylation sites (11,14). While identified ncRNA are mostly involved in heart-related processes (miR-1, miR-133a, miR-208b, MIAT), DNA methylation mainly occurs in genes involved in immune response, nervous system and ion transport. The binding of some transcription factors, including PAX5, RUNX3, SPI1, IRF4 and TBX21, was predicted to be altered by a differential methylation in targets promoter region, supporting the impact of methylation in this disease.

Although these studies offer an initial analysis of the factors contributing to the development of CCC, they are constrained to the combination of only two types of omics data, providing only a partial perspective on the underlying mechanisms. To get a more complete comprehension of the pathogenic process occurring in CCC myocardium, we performed here a multi-omics integration of transcriptomic, miRnomic, methylomic and mitochondria sequencing data.

Materials and methods

Raw data collection

All raw data comes from our previous studies, including bulk RNA-sequencing (11), miRNA sequencing (12), Illumina EPIC methylation (11) and mitochondria variant calling. Human left ventricular free wall heart tissue samples were obtained from patients with end-stage heart failure CCC at the time of heart transplantation (n=7). CCC patients underwent a serological diagnosis of *T. cruzi* infection and standard electrocardiography and echocardiography, and tissues were subject to

histopathological assessment as previously described. The protocol was approved by the institutional review boards of the University of São Paulo School of Medicine and INSERM (French National Institute of Health and Medical Research). Written informed consent was obtained from all patients. All experimental methods comply with the Helsinki Declaration. Each dataset was aligned on the GRCh37 (hg19) human reference genome from Ensembl, and all classical pre-processing steps have been applied as described in previous study (11,12).

| CTRL | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|----------------------|---|---|---|---|---|---|---|
| Transcriptomic | x | x | x | x | x | x | x |
| Methylomic | x | x | x | x | x | x | x |
| miRnomic | | x | x | x | x | x | x |
| Mitochondria genomic | x | x | x | x | x | x | x |
| CCC | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Transcriptomic | | x | x | | | x | |
| Methylomic | x | x | x | x | x | x | x |
| miRnomic | x | x | x | x | | x | |
| Mitochondria genomic | x | x | x | x | x | x | x |

Table 1. Human heart tissue available through the different omics.

Missing samples imputation

Across the multi omic dataset, 6 samples are missing in at least one omic: 4 in transcriptomic and 3 in miRNomic (**table 1**). To complete the multi-omics dataset, missing values for samples were imputed using the KNN-imputer function of scikit-learn python library. All features' values were centered and scaled to be comparable, and a custom matrix distance for the multi-omics dataset was made, considering for each omic: the distance between all samples, the number of features and the number of missing samples. The imputation of controls and cases was done independently. For each phenotype, the following steps have been performed:

Step 1. For each omic M and for each couple of samples (i,j) in M, Euclidean distance metrics has been computed:

$$\forall M, \forall_{i,j} d_M(v_i, v_j)$$

Step 2. For each distance matrix obtained from step 1, each distance value between a couple of samples (i,j) was divided by the number of features N_M in omic M. Since the number of features is very different from one omic to another, this step allows to normalize the distance between two samples in each omic by the number of features:

$$\forall M, d'_M(v_i, v_j) = \frac{1}{N_M} d_M(v_i, v_j)$$

Step 3. An average distance matrix has been computed. The normalization term of the mean corresponds to the number of available samples through omics:

$$D(v_i, v_j) = \frac{1}{P} \sum_{M=1}^P d'_M(v_i, v_j)$$

Step 4. The missing samples imputation is performed with the KNN imputer function of sklearn python 3 package (v.1.2.2), using the custom distance matrix made previously.

By this way, both differences in omic size and number of missing samples are normalized in the final distance matrix, and all omics have the same contribution in data imputation.

Data augmentation

An estimation of the optimal number of samples to ensure a statistical power greater than 0.8 for the multi-omics integration was performed using MultiPower (v.1.0). To reach the desired numbers, synthetic samples were created as linear combinations of randomly selected original samples of the same phenotype. For each synthetic sample, random weight between 0 and 1 were drawn for each selected sample so the sum of all weight is reach 1. To ensure synthetic sample remain in a region of the samples space that can have a biological meaning the first weight, called α , was chosen to be between 0.8 and 1. If we note $\{P_i\}_{i=1,\dots,K}$ a sampling of K samples in the N original samples and P' the new synthetic sample we have:

With $\alpha_1 \in [0.8, 1[$ and $\alpha_i \in]0,1[$
for $i = 2, \dots, K$ such that $\sum_1^K \alpha_K = 1$

$$P' = \sum_1^K \alpha_K P_K$$

In total, 193 synthetic samples were generated per phenotype, to obtain a total cohort of 400 samples.

Feature reduction

Given the large number of features in some omics, a feature selection was applied to both transcriptomic and methylomic datasets. For the transcriptomic features, a filtering process was applied based on gene function, retaining only transcription factors, protein-coding genes, and long non-coding RNA (lncRNA). A second feature selection was performed in lncRNA, keeping only those annotated in lncRNADisease (15). In the case of methylation sites (CpG), due to their large quantity, only CpGs located in regulatory regions of genes (TSS-1500 and 5'UTR region) included in ChagasDB (9) were included in the analysis.

Multi-omic integration

The multi-omic integration was conducted using MOFA+ (v.1.8.0) (16) with the selected features. All the following analysis was performed using R (v.4.2.1). For transcriptomic expression data (RNA sequencing and miRNA sequencing), a log2 transformation was applied to ensure a Gaussian distribution. CpG methylation beta values were transformed into M-values (17) and modeled using a Gaussian likelihood, similar to the approach employed in the original MOFA+ analysis (16). The variant allele frequency of mitochondria was binarized, representing the presence or absence of the mutation, and modeled using a Bernoulli distribution. The analysis was iterated 19 times, gradually increasing the number of factors from 2 to 20. To ensure the selection of meaningful factors, only those representing at least 1% of the data variability within a single omic were retained. The specific function of each factor was evaluated using MOFA GSEA enrichment on Reactome database (18). For each factor of interest, the weights of features within that factor were scaled individually

by omic. Only the top features with sum of weights equal to or greater than 25% of the total feature weights within that specific factor were retained. A similar analysis was performed using CCC patients only. In this case, the features weight cut-off was fixed to 50% of the total feature weights.

Pre-processing pipeline validation

The quality of the data augmentation was tested using 1643 samples from TCGA (19), including breast (BRCA n = 1106) and lung cancer (LUAD, n = 537) transcriptomic data (**supplementary table 1**). For each dataset, different oversampling methods were applied, using each time 10 existing samples to generate 200 synthetic samples. The other methods include bootstrapping, SMOTE (v.0.4.1) (20) and KNNOR (v.0.0.4) (21). The validity of the results obtained with MOFA was assessed by comparing them with those obtained with 3 other multi-omics integration methods: DIABLO (MixOmics) (v.6.24.0) (22), JIVE (v.2.4) (23) and MCIA (v.1.40.0) (24). Using the same filters as before, the obtained features were compared using the Jaccard index, and the overlap coefficient. Finally, the complete analysis was performed 10 times. For each run, one raw sample was randomly removed in one omics. The results were compared using the Jaccard index, and the minimal overlap.

Multi-omics correlation network

In order to understand the link between the selected multi-omics features, two correlations networks were made, specific for each phenotype (healthy controls and CCC). The transcriptomic data was split into 3 different types of features: protein coding genes, transcription factor and lncRNA. Correlation clusters were assessed using the pheatmap R package (v.1.0.12) distinct biological functions associated with each cluster were identified using the R gprofiler2 package (v.0.2.1). To ensure the reliability of the identified correlation, the network were filtered using interaction public databases, including STRING (25), BioGRID (26), Coexpedia (27), GRNdb (28), IntAct (29), lncTarD (30), miRecords (31), miRTarBase (32) and TarBase (33).

Lymphocyte T and B repertoire analysis

An analysis of immune repertoire was performed with TRUST4 (v.1.0.8) tool on tissue RNA-seq samples (7 control, 10 severe CCC and 8 dilated cardiomyopathy) (34). TRUST4 align reads against V, D, J and C genes, and exports all the distinct clonotypes found in each sample. These clonotypes were then analyzed thanks to the ImmuneArch R package (v.0.8.0), in order to identify shared clonotypes and gene usages between samples of the same phenotype (35). Clonotypes sequences were clustering using Clustal Omega (v.1.2.4) (36).

Results

Validation of the pre-processing pipeline

Multi-omics analysis was conducted on a total of 14 human heart tissue samples, comprising 7 healthy controls and 7 CCC cases, and 4 different omics (transcriptomic, methylomic, miRnomic and mitochondria sequencing). Based on multi-power tool estimation, the statistical power using so few samples fall within the range of 0.1 to 0.3, depending on the specific omics under evaluation. According to this estimation, it was necessary to increase the

sample size to a minimum of 75 samples per phenotype to achieve a minimal statistical power of 0.8 in each omic (**supplementary figure 1A**). Therefore, synthetic samples were generated using the barycenter metrics, as detailed in the materials and methods section. To confirm the stability of the analysis result, synthetic datasets of increasing sizes (from 21 to 400 samples per phenotype) were created and submitted to a multi-omics integration. When the select features after multi-omics integration with the largest cohort (n = 400) were compared with varying lower sample counts, it was observed that the same features were consistently found for 100 samples and more (**supplementary figure 1B**). Given the high number of features in different omics, we opted to double the sample size to 200 per phenotype, aiming to achieve a statistical power of almost 1 for each omic, as a statistical power of 0.8 might prove insufficient. Compared to other existing methods, the barycenter approach better preserves the raw data distribution observed in PCA compared to bootstrapping and KNNOR and is similar to SMOTE (**supplementary figures 2**). About the multi-omics integration itself, it has been performed using the MOFA+ methods, which

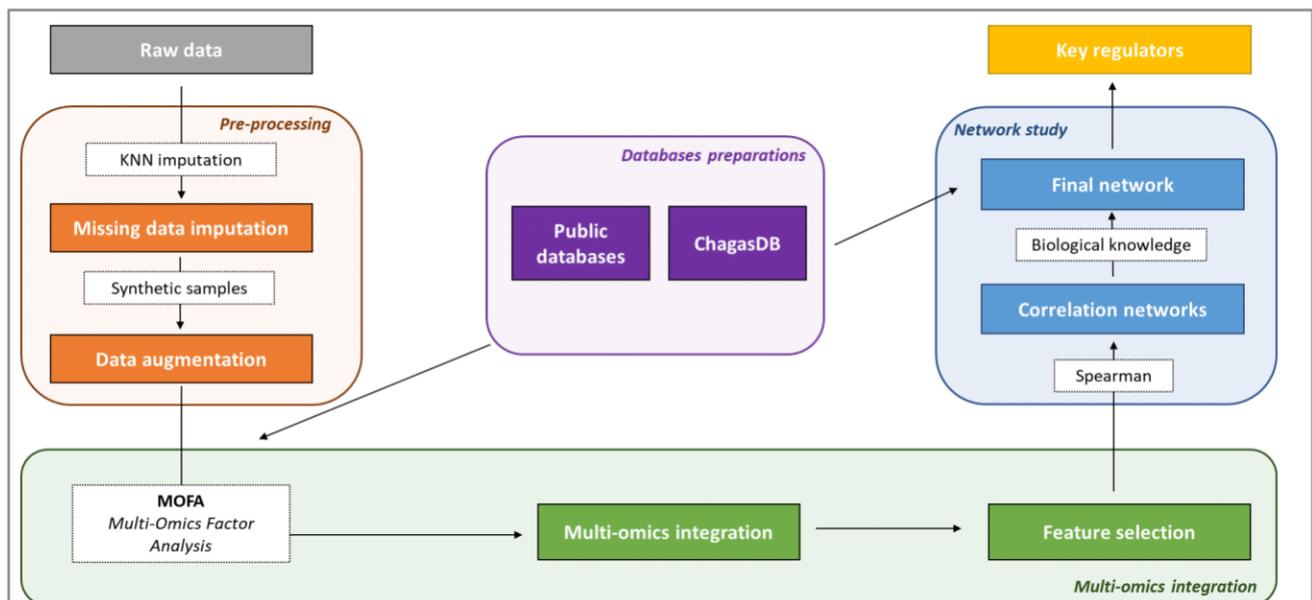


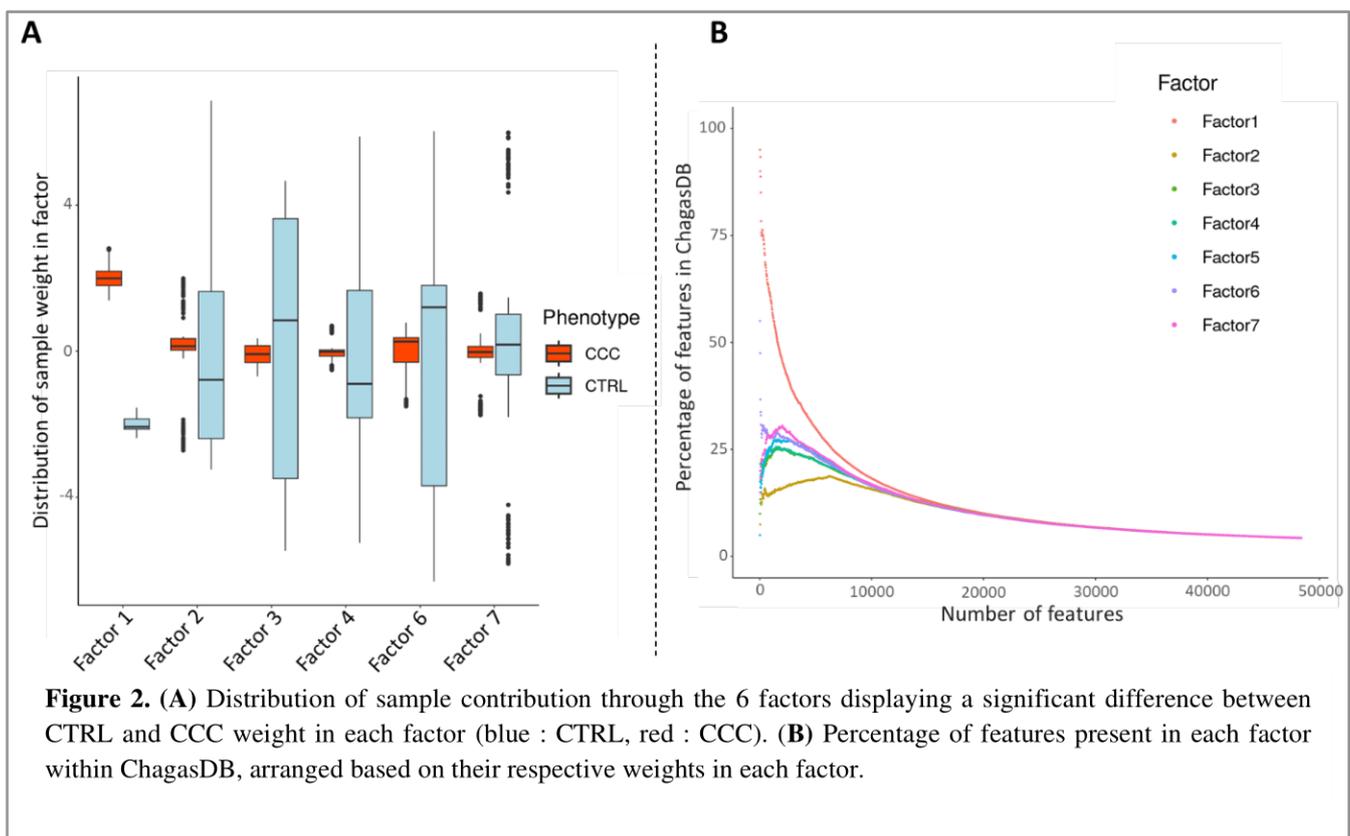
Figure 1. Workflow overview. Raw data were obtained on end-stage CCC heart tissue and healthy heart tissue. All the individual omics pre-processing, from alignment to normalization, were performed as described in previous study. Multi-omics pre-processing steps (orange) include missing samples imputation with KNN-imputer from scikit-learn library, and synthetic samples generation using barycenter approach. The multi-omics integration itself was performed using MOFA+ tool (green). The features selected at multi-omics level were used to create phenotype-specific correlations network (blue). The downstream analysis was made on correlation network filtered with known or predicted interactions in public databases.

provides similar results to JIVE and MCIA (**supplementary figure 3A**), with 40% (1361/3338) of MOFA+ features in common with JIVE or MCIA. Particularly, 70% of MOFA+ features were retrieved in MixOmics integration, 96% in JIVE and 41% in MCIA (**supplementary figure 3B**). Those overlap, considering the vast number of original features, strongly indicate that the selected features were not randomly chosen and hold relevance for studying CCC. To validate the effectiveness of the overall pipeline, we repeated the process 10 times using different generated datasets, each time excluding one existing original sample. By this way, the values of this sample will be imputed, and some synthetic samples will be generated from his imputed values. All iterations produced consistent outcomes, with a minimum Jaccard index of 70% and a minimum overlap coefficient of 90% in shared features (**supplementary figure 4**). These consistent results strongly support the robustness of this pipeline (**figure 1**).

Factor 1 is associated with Chagas disease Cardiomyopathy

The MOFA integration was conducted on a subset of the complete dataset, consisting of 266,256 features across the 4 omics (transcriptomic n = 19,559,

methylomic n = 246,660, miRnomic n = 652, mitochondria sequencing n = 45). According to MOFA integration, 7 factors carry at least 1% of the data variability. The 2 first factors carry the most variance in different omics. Whereas factor 1 is strongly supported by transcriptomic and methylomic data, factor 2 is carried by mitochondria sequencing (**supplementary figure 5A**). No correlation was found between pairs of factors, indicating each of them is carrying specific information (**supplementary figure 5B**). Whereas 6 factors display a significant difference between CTRL and CCC weight (Wilcoxon p-value ≤ 0.05), the factor 1 is the most discriminative between CCC and healthy controls (**figure 2A**). Moreover, most of the top-rank genes in factor 1 are known to be involved in Chagas disease cardiomyopathy, which is less the case for other factors (**figure 2B**). Interestingly, some of the top-ranked features are absent in ChagasDB, and therefore, remain unknown in the context of this disease. Considering the interest of factor 1 in this disease, a GSEA enrichment was performed using all the features rank in this factor, using the Reactome database. The positive features, so the features positively enrich in CCC, are involved in both innate and adaptative immunity (Complement cascade, TCR/BCR



signaling...) (**supplementary figure 6**). Interestingly, the interleukin-10 signaling, associated to anti-inflammatory response, is also enriched. On the opposite, some cardiac and mitochondrial processes are negatively enriched. All these results are in line with the current knowledge of CCC. Considering the importance of factor 1 on the disease, the top-features associated to this factor were selected, including 1,214 genes, 2,066 methylation site and 58 miRNA. The majority of the selected genes are involved in immune response, mostly associated to T cell and inflammatory response (**supplementary table 2**). However, most of top-ranked positive features are known to be involved in B-cell related process (FCRL5, IGJ, CLNK, FAM30A, NUGGC, CD79A, IGLL5), suggesting an important role of this cell in CCC.

CCC patients presents specific regulatory patterns

To identify deregulated processes in CCC, two correlation networks were constructed, one based uniquely on healthy control samples, and the other only on CCC samples. Only correlations between

expressed features were retained, with methylation sites assigned to the associated gene. To investigate the potential involvement of regulatory elements, the transcriptomic dataset was divided into three types of features: protein coding, transcription factors (TF), and long non coding RNA (lncRNA). Interestingly, the correlation levels between some pairs varies depending on the phenotype (**supplementary figure 7A**). Given these observations, the correlation networks were filtered as follows: the absolute correlation level must be over 0.8 in one phenotype, and less than 0.4 in the other (**supplementary table 3, supplementary table 4**). Although the same features are present in both networks, their behavior is significantly different between CCC and CTRL (hypergeometric p-value $\leq 2E-16$) (**supplementary figure 7B/7C**). Looking at the different impact of regulatory features on gene expression, we perform a hierarchical clustering (HCA) to group them according to their correlation with protein-coding genes. 3 specific clusters appear, illustrating clear different co-expression patterns (**figure 3A**) (**supplementary table 5, supplementary table 6**).

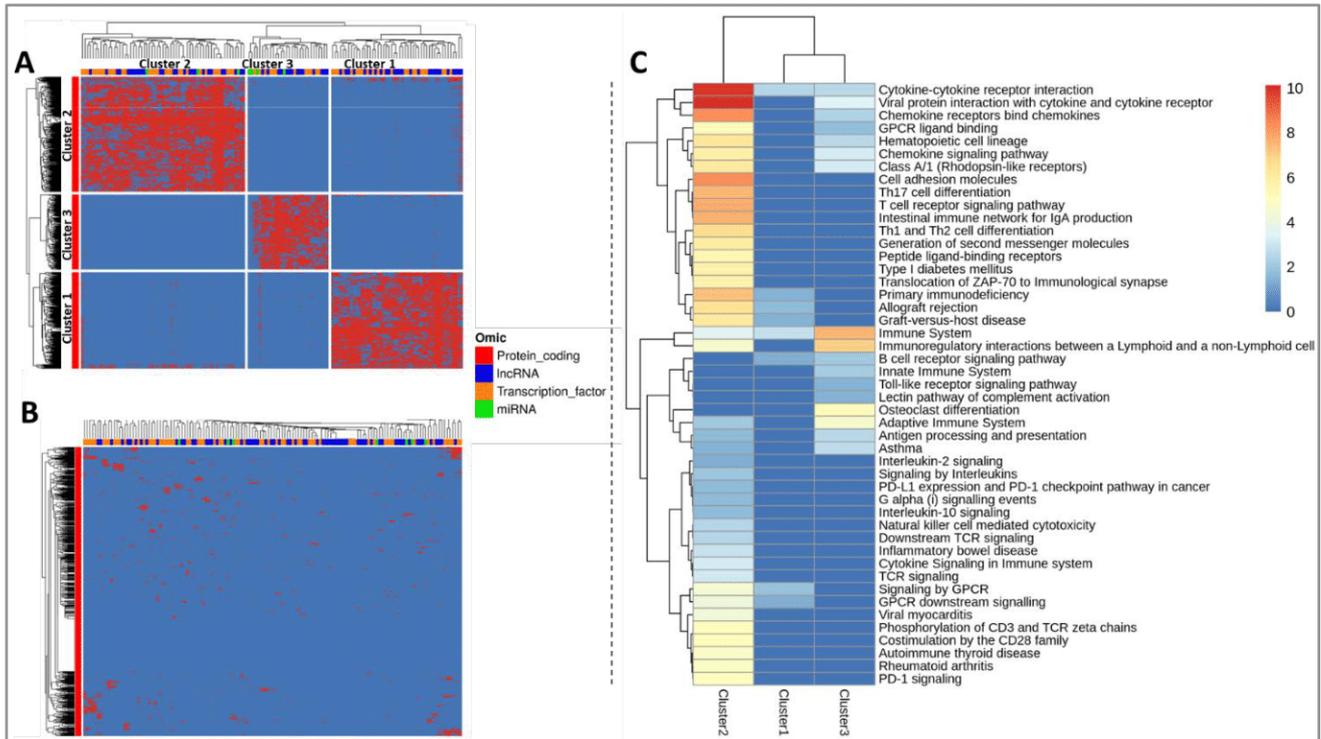


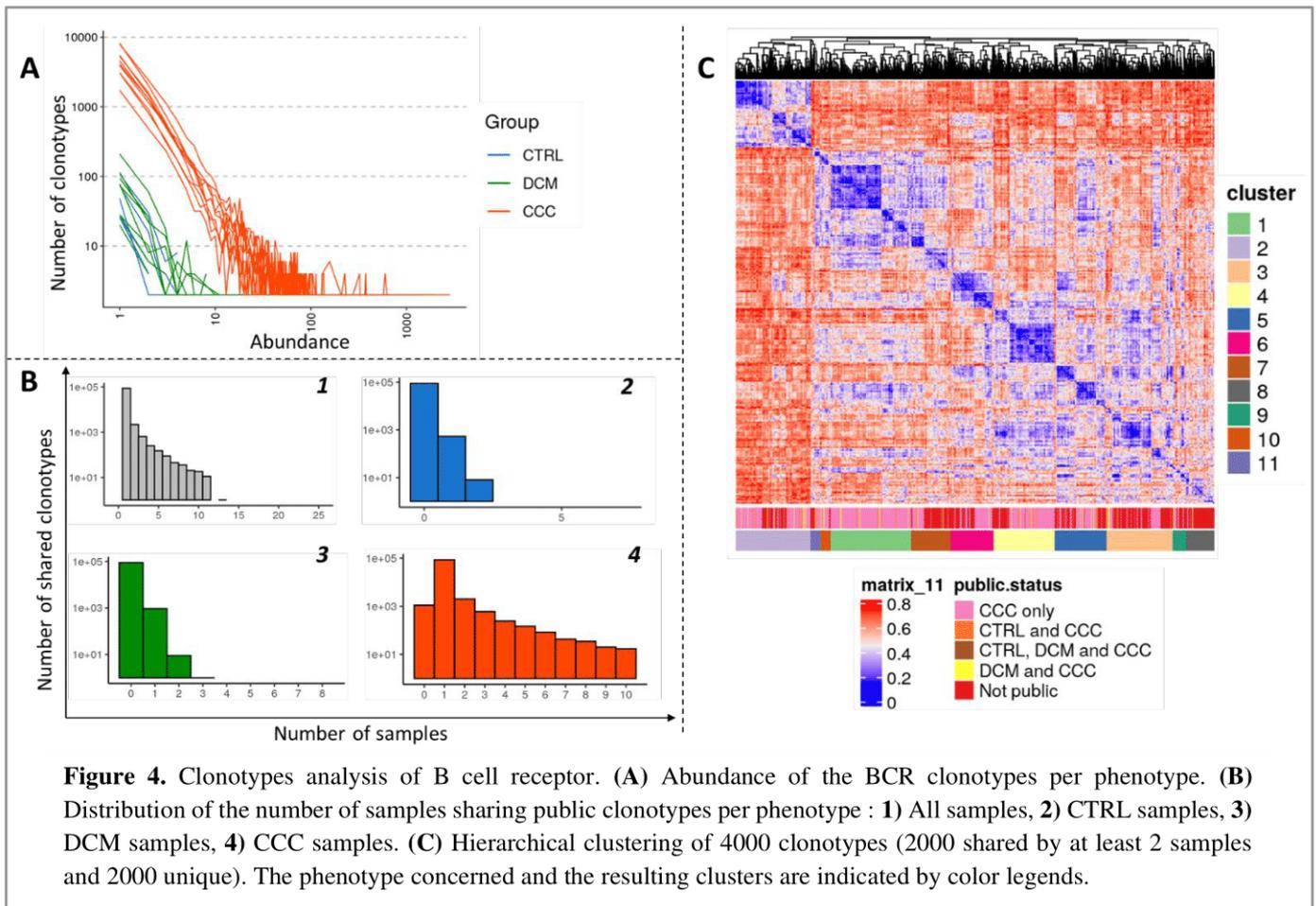
Figure 3. (A) Hierarchical clustering of the regulatory features (lncRNA, miRNA, transcription factor) correlated with protein coding genes in CCC network. **(B)** Clustering of the regulatory features (lncRNA, miRNA, transcription factor) correlated with protein coding genes in CTRL network. **(C)** KEGG and Reactome pathway significantly enriched in at least one of the CCC network cluster.

This phenomenon is not observed in healthy controls (**figure 3B**). Using KEGG and Reactome database, we identified 47 pathways significantly enriched in at least one cluster (**figure 3C**). Features in cluster 2 are associated to T-cell related processes (T cell receptor signaling pathway, Th1 and Th2 cell differentiation, Th17 cell differentiation), whereas cluster 3 is associated to B cell response (B cell receptor signaling pathway) and innate immune response (Toll-Like Receptor, complement activation). None of the KEGG or Reactome pathways were specific to the cluster 3. Interestingly, pathways related to anti-inflammatory response, including IL-10 and PD-1 signaling affiliated to cluster 2. Moreover, some autoimmune diseases (type 1 diabete, inflammatory bowel disease...) were also associated to cluster 2, as well as viral myocarditis.

B cell receptors present a high convergence in all CCC

Considering the previous results highlighting the importance of B and T cell-receptor signaling in CCC, an analysis of BCR and TCR repertoire was performed. The TCR repertoire analysis uncovered that the diversity and abundance of clonotypes were higher in CCC than in DCM or CTRL (**supplementary figure 8A**). However, no clonotype was found to be shared across all CCC (**supplementary figure 8B**).

Analysis of the BCR heavy and light chains revealed that the diversity and abundance of clonotypes was much greater in the CCC samples than in the CTRL or DCM samples (**figure 4A**). This result is a first control that confirms an increased presence and/or activity of B cells in the samples, in association with chronic inflammation. The specific study of light chain clonotypes by considering the CDR3 sequence and the corresponding V/J genes reveals that many clonotypes are common to several CCC samples and that more than 10 of them are common to all 10 samples (**figure 4B**). This phenomenon is not found



in the CTRL or DCM samples in which the clonotypes are shared by a maximum of 2 samples. The detection of this public clonotype repertoire seems to indicate the presence of sequence convergence by specific selection, similar to that found in the presence of specific antigens. However, the absence or very low parasitemia in the patients' cardiac tissue seems to rule out the possibility of the presence of pathogen antigens. In order to be more precise, we wanted to check whether this convergence was found in a larger proportion of clonotypes and whether similar but not identical sequences to the public repertoire could be detected. We thus selected 2000 clonotypes shared by at least 2 samples and having the longest CDR3 sequences. As a control, we then selected 2000 clonotypes present in a single sample and having the longest CDR3 sequences. Using Clustal Omega, we aligned the sequences and then calculated the distances between these 4000 sequences (**figure 4C**). The distance matrix clearly shows the convergence of clonotype series around different clusters. We used an HCA algorithm to cluster the sequences by their proximity. 11 clusters were thus identified (with an arbitrary choice of cutting the hierarchical tree). For each cluster, we studied the frequency distribution of the V genes associated with the sequences. The majority of clusters are dominated by the presence of one or two V genes, often close to each other (**supplementary figure 9**). Those clusters are also dominated by clonotypes of the public repertoire. With regard to the BCR heavy chains, while a strong clonal expansion was always observed in the CCC samples but not in the CTRL or DCM samples, the presence of a public repertoire was not observed. The stoichiometric relationship between light and heavy chains remains close to the expected value (expected value: 1.5, observed value: 1.17). As the heavy chain is commonly more mutated than the light chain, it is possible that the sequencing depth we used was not sufficient to reveal the commonality of certain sequences, given the very large possible repertoire.

Dual and opposite role of lymphocytes-related transcription factor

To identify potential regulatory interactions among the correlated features, we compared our correlation networks with various interaction databases (see Methods). Both healthy control and CCC specific correlations are significantly enriched in known interaction (CTRL hypergeometric test p-value p-value < 3.68E-249, CCC hypergeometric test p-value < 2.87E-08), confirming their relevance. Consequently, we filtered the correlation networks, keeping only correlated features known or predicted to interact together. With the aim of characterizing the mechanisms that contribute to the exacerbation of the inflammatory response in CCC, we focused on the correlations/interactions between different type of features previously assigned to cluster 2, mainly involved in T cell activation and TCR signaling. As expected, the most interconnected features are protein coding genes and transcription factors, whereas miRNAs and lncRNAs have a limited number of targets. For example, RP11-357H14-17, the top-correlated lncRNA assigned to cluster 2, was removed from the interaction network. Looking at the 10-top connected TF, all of them are related to T- or B cell activation and development. Moreover, they seem to combine to regulate distinct sets of genes (**supplementary figure 10A**).

We have identified 9 TF that exhibit correlations with TCR genes, suggesting activation through the TCR signaling pathway. Notably, 7 of them also show correlations with TLR and/or BCR genes. Furthermore, they are associated with genes participating in 9 specific biological functions of interest (**figure 5A**). IRF4, IKZF1, and PAX5 are linked to all these processes. IRF4 displays an inverse correlation with certain cardiac-related genes, such as MYH6 and RRAD (**figure 5B**). On the opposite, it exhibits a positive correlation with genes involved in inflammation regulation, as well as other lymphocytes-related TF (TBX21, PRDM1, BATF, PAX5, IKZF1) (**figure 5C**). Among the genes related to this process, RHOH is the only gene that is downregulated in CCC. MIR551b demonstrates an inverse correlation with SOCS3, suggesting a potential involvement of this miRNA in the regulation of inflammation.

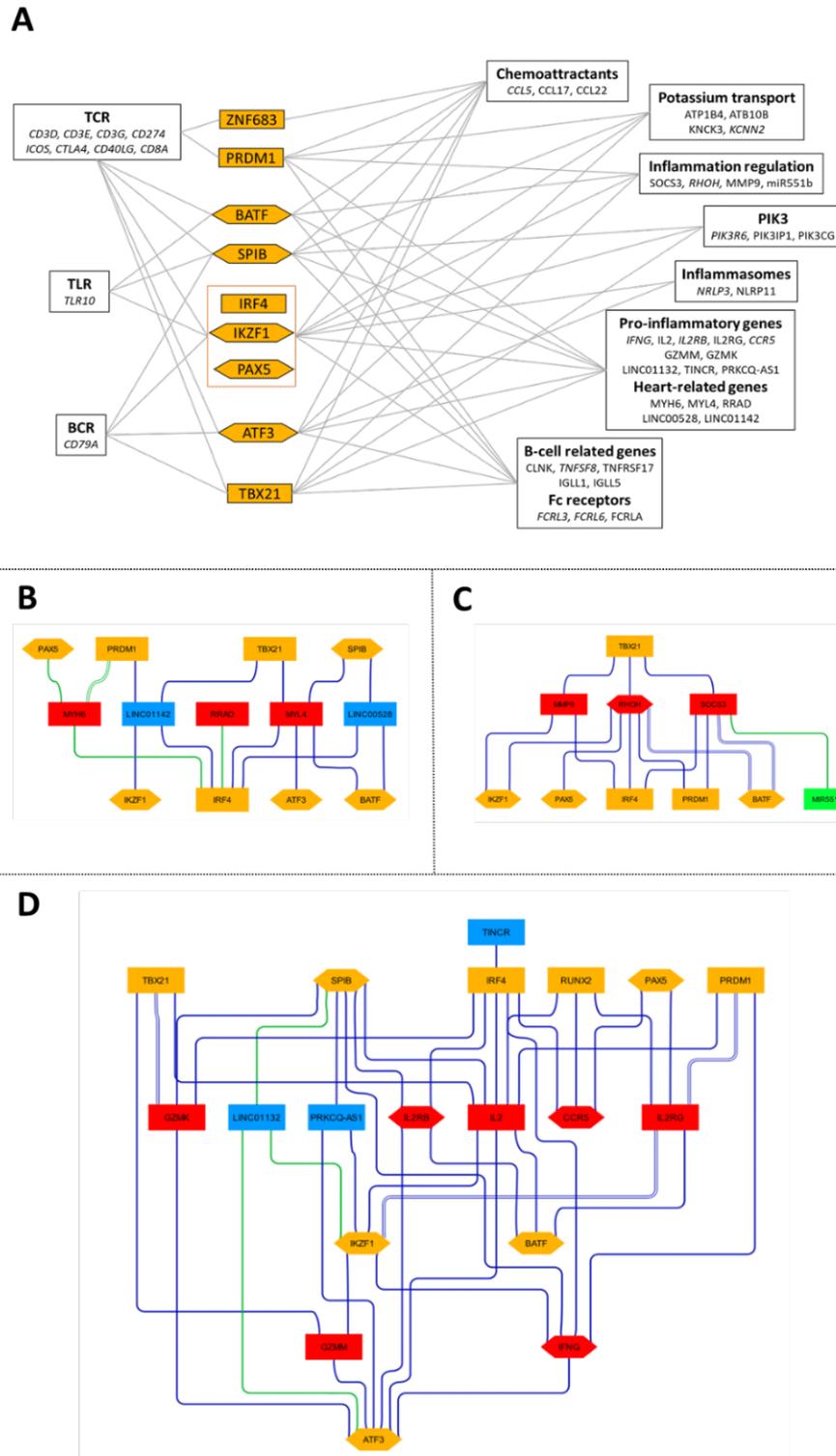


Figure 5. (A) Diagram illustrating the impact of top transcription factors on the biological processes influenced by genes within cluster 2. (B) Correlation network of the transcription factor targets genes involved in heart-related processes. (C) Correlation network of the transcription factor targets genes involved in inflammation regulation. (D) Correlation network of the transcription factor targets genes involved in inflammation. Nodes color: red: protein-coding genes, blue: lncRNA, green: miRNA, orange: transcription factor. Nodes shape: square: feature expression, hexagon : feature expression carrying at least one CpG selected with MOFA+ in its promoter. Edge line color: blue: positively correlated, green: negatively correlated. Edges line type: line: correlated in CCC network, dot: correlated in CTRL network, double line: correlated in both networks.

On the other hand, IRF4, along with nearly all other transcription factors, shows a positive correlation with pro-inflammatory genes. LINC01132, a down-regulated long non-coding RNA implicated in NF- κ B signaling, exhibits a negative correlation with ATF3, SPIB and IKZF1. Interestingly, both of genes and TF presents sometimes at least one differentially methylated CpG in their promoter region. Along with miRNA and lncRNA, those epigenetic marks may explain the involvement of same TF in opposite pathways (pro and anti-inflammatory processes). This dualistic behavior could also be explained in variations among the TF combinations. To illustrate, IRF4, SPIB, and ATF3, which are among the most interconnected transcription factors within the entire network, demonstrate distinctive co-regulated gene clusters either independently or in concert (**supplementary figure 11**).

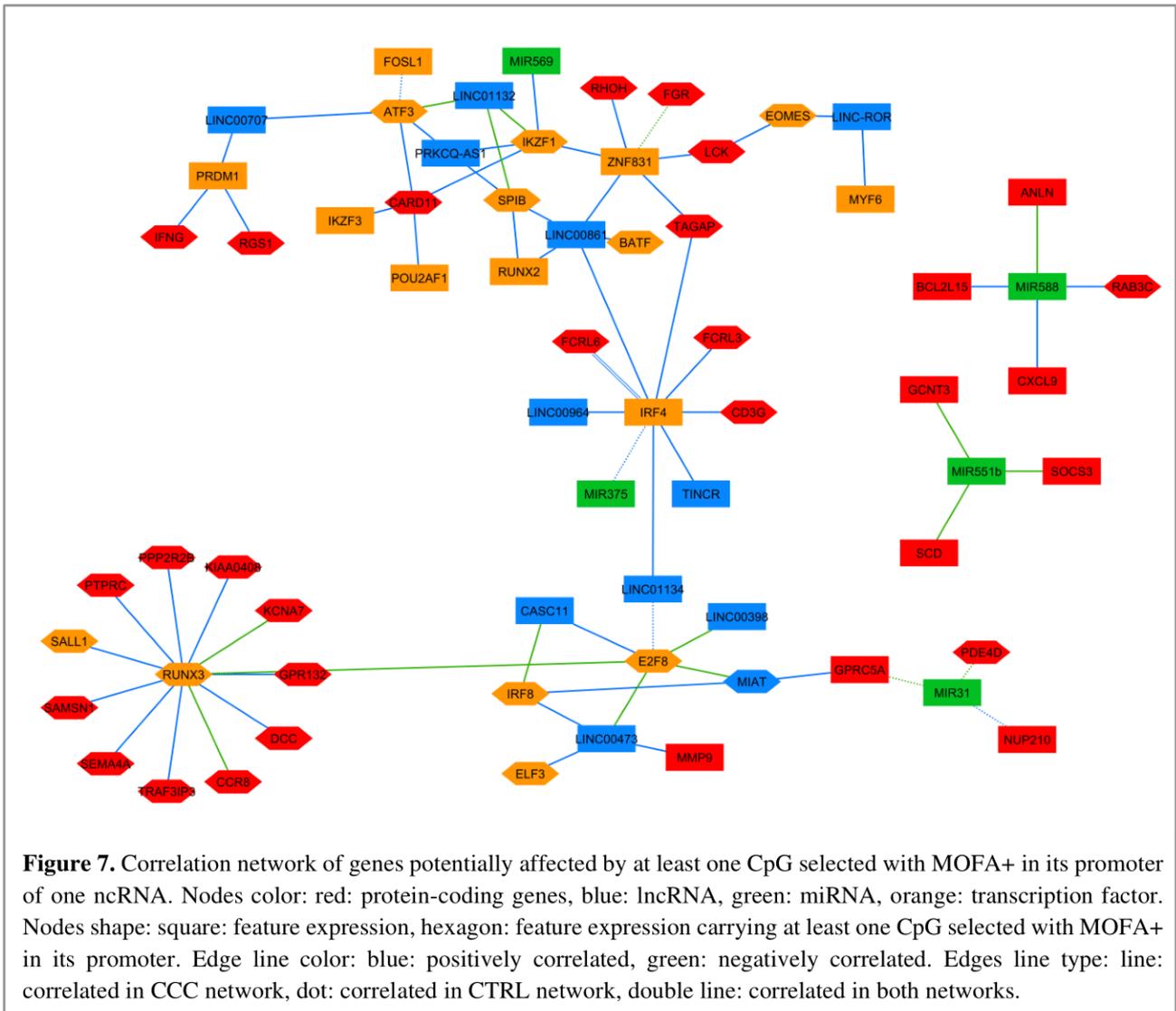
TLR may regulate innate immune response

Considering the high convergence observed in BCR, suggesting a high-pressure selection on clonotype and therefore a continuous stimulation, we performed the same analysis on cluster 1 and 3, significantly enriched in BCR and TLR signaling. SPIC is the most highly connected transcription factor and appears to independently target numerous genes. To a lesser extent, it may also cooperate with FOSL1, a transcription factor involved in the regulation of inflammation (**supplementary figure 10B**). Focusing on potential origin of inflammation response, we have identified several genes involved in both innate (TLR, complement) and adaptive (BCR, TCR) immune system. 7 transcription factors showed correlations with at least one of these processes, with 6 of them associated with BCR genes. All those TF are correlated with inflammatory response. Interestingly, RUNX3, IRF8, SPIC and MYB are also correlated with immune regulation related genes (**figure 6A**). Looking at the resulting networks, SPIC is negatively correlated with TLR6, TLR8 as well as inflammasome (PYCARD) and mitochondria (CYP4F22) genes, but positively correlated with TLR7 (**figure 6B-6D**). While SPIC is negatively correlated with GAPT, LAIR1 and TNFAIP8L2, all associated to adaptive immune response regulation, it demonstrates positive correlations with PLD4 and GAPT, two genes linked

to the regulation of innate immune responses (**figure 6E**). Similarly, it also exhibits a negative correlation with numerous genes linked to inflammation, including LINC01010, a target of MyD88 (**figure 6F**). In a lesser extent, ASCL2 and MYB, correlated with C2 gene, are also negatively correlated with some inflammatory genes. Interestingly, ASCL2 is negatively correlated with CCDC88B, required for T-cell pathway, in control network. Considering the DNA methylation observed in both ASCL2 and CCDC88B promoter, we can hypothesize that methylation may influence the regulation of CCDC88B by ASCL2. All these results tend to demonstrate the involvement of SPIC in innate immune regulation, probably through TLR7 stimulation.

Epigenetic regulation at the heart of CCC pathogenic process

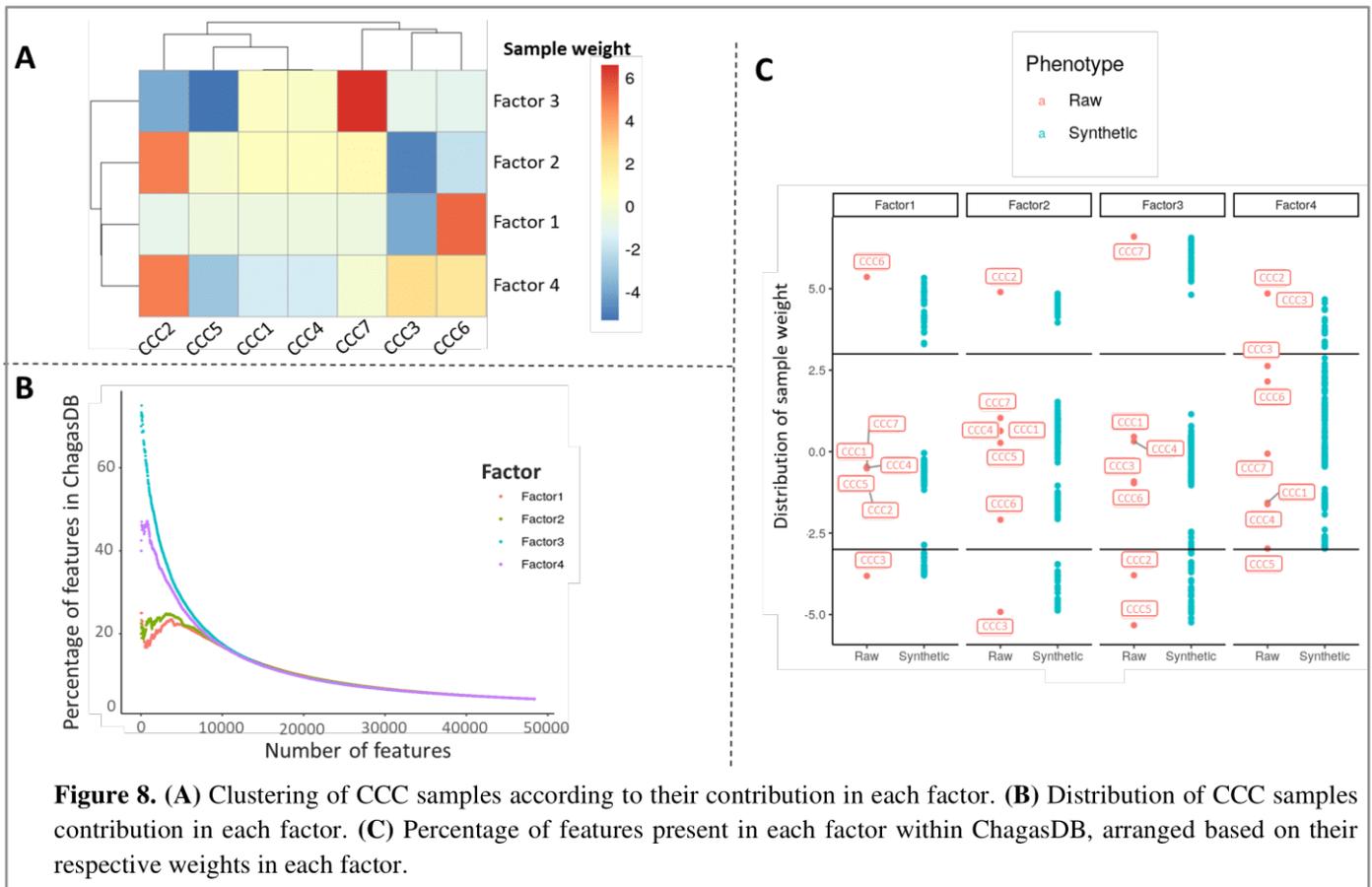
Considering the high number of differentially methylated genes and transcription factors observed in cluster 1, 2 and 3, a stringent filter was applied to the network, selecting only protein coding and transcription factors correlated at both expression and methylation level in one phenotype, or inversely correlated in both phenotypes. By this way, we ensure that the methylation may be responsible of the differential expression observed. As lncRNA and miRNA may not be affected by methylation in gene promoter, we conserved all these correlations. The resulting network highlight the potential impact of RUNX3 on pro-inflammatory genes expression, as well as Th2 T-cell repression through CCD8 (**figure 7**). Moreover, some miRNA may have a crucial role in the exacerbate inflammation occurring in CCC, such as MIR588, positively correlated with BCL2L15, RAB3C and CXCL9. On the opposite, the lack of miRNA binding may enhance inflammatory process, like MIR375, known to reduce apoptosis in cardiac cell, and correlated with IRF4 in control network only. Given the high number of lncRNA correlated with IRF4, it is plausible that its activity is sequestered or modulated by one of them. It is notable that IFNG also carried at least one differentially methylated CpG in its promoter and may be up regulated by PRDM1 transcription factor.



While this analysis did not reveal many TF-TF potential interactions (based on both expression and methylation correlations), many TF selected by MOFA+ have at least one CpG differentially methylated in their promoter. The methylation level may not gradually affect the binding but act as a switch. Looking on all the TF-TF correlations, we observed that PAX5, IRF4 and ATF3 are the top-connected TF, but IRF4 is the only one that do not carry any differentially methylated CpG site (**supplementary figure 12**). Although alterations in methylation may not definitively signify whether the correlated transcription factor will or will not bind to this promoter in CCC, it remains interesting to explore this possibility for a deeper understanding of the pathogenic process.

MOFA factor discriminate 5 of the 7 CCC patients

Beyond the shared variability observed across all CCC patients compared to healthy controls, a secondary multi-omics integration was conducted with the purpose of examining individual patient differences, with the aim of determining whether the disease may have diverse origins. 4 factors carrying at least 1% of the variability in the data were used to discriminate 5 of the 7 CCC patients (**figure 8A**). Features with positive or negative weights in these factors are enriched in Reactome pathways related to mitochondrial function (respiratory electron transport, complex I biogenesis), innate immune response (complement cascade, Toll Like Receptor Cascade, interferon signaling), nervous system or muscle contraction (**supplementary figure 13**). As all these processes are directly linked to chronic



| | CCC6 | | CCC3 | | CCC2 | | CCC7 | | CCC5 | |
|-----------------------|-------|--------|-------|--------|-------|--------|-------|--------|-------|--------|
| | Total | Unique |
| Protein coding | 1309 | 585 | 3045 | 1344 | 1489 | 73 | 1546 | 5 | 2720 | 36 |
| Transcription factor | 84 | 28 | 228 | 112 | 87 | 3 | 124 | 1 | 191 | 4 |
| lncRNA | 86 | 18 | 312 | 158 | 143 | 4 | 134 | 0 | 244 | 2 |
| miRNA | 57 | 10 | 102 | 24 | 120 | 38 | 77 | 1 | 95 | 32 |
| Methylation site | 3093 | 1179 | 5815 | 2430 | 6395 | 2207 | 3280 | 1301 | 6420 | 3144 |
| Mitochondria mutation | 11 | 9 | 5 | 2 | 12 | 0 | 9 | 0 | 14 | 13 |

Table 2. Top 50% features selected in each factor and associated to each sample. The total number of features, and the number of not shared features across the samples are displayed.

chagasic cardiomyopathies, the discriminative features obtain through the different factors may be involved in the pathogenesis. Furthermore, except for factor 3, the top-features of the factors are only partially known in ChagasDB (**figure 8B**).

CCC patients present specific expression profiles

In order to understand the specificity of each patient, the 50% of features that contributed most to the

various factors were selected (**supplementary table 7**). They were then associated with each sample according to their contribution to each factor. Considering the synthetic samples associated to each raw samples, 5 clusters of samples were selected, assigned to CCC6 (n=20), CCC3 (n=19), CCC2 (n=19), CCC7 (n=20) and CCC5 (n=17) (**figure 8C**).

The 5 sample clusters bring together genomic, transcriptomic and methylomic data. However, with the exception of methylation sites, few of these are sample-specific (**table 2**). This is particularly the case for the CCC2 and CCC7 clusters, where a small number of protein-coding genes have been selected. Three samples stand out from the rest with specific features involved in precise biological processes: CCC6 (mitochondrial alterations, notably complex I), CCC3 (activation of T-CD4 lymphocytes and the Wnt pathway) and CCC5 (Toll-like receptor signaling activity) (**figure 9**). As CCC3 and CCC6 are the only samples not imputed characterized by specific biological processes, further analysis was conducted only on these samples.

Mitochondria mutation as a cofactor of CCC pathogenesis

Interestingly, 9 mitochondrial mutations were specifically assigned to CCC6 cluster. Among them, MT-ND1 3604 CTAG > C was the top-weighted mutation in factor 1 (weight = 3.43) compared to others (weight < 1.55). Regarding the differential gene expression between the samples carrying this mutation and the other, 2 miRNAs present significant differential expression ($abs(log2FC) \geq 1.5$ & adjusted p-value ≤ 0.05): miR-369-5p and miR-591. Therefore, this mutation may be involved in the difference of expression observed with these 2 miRNAs.

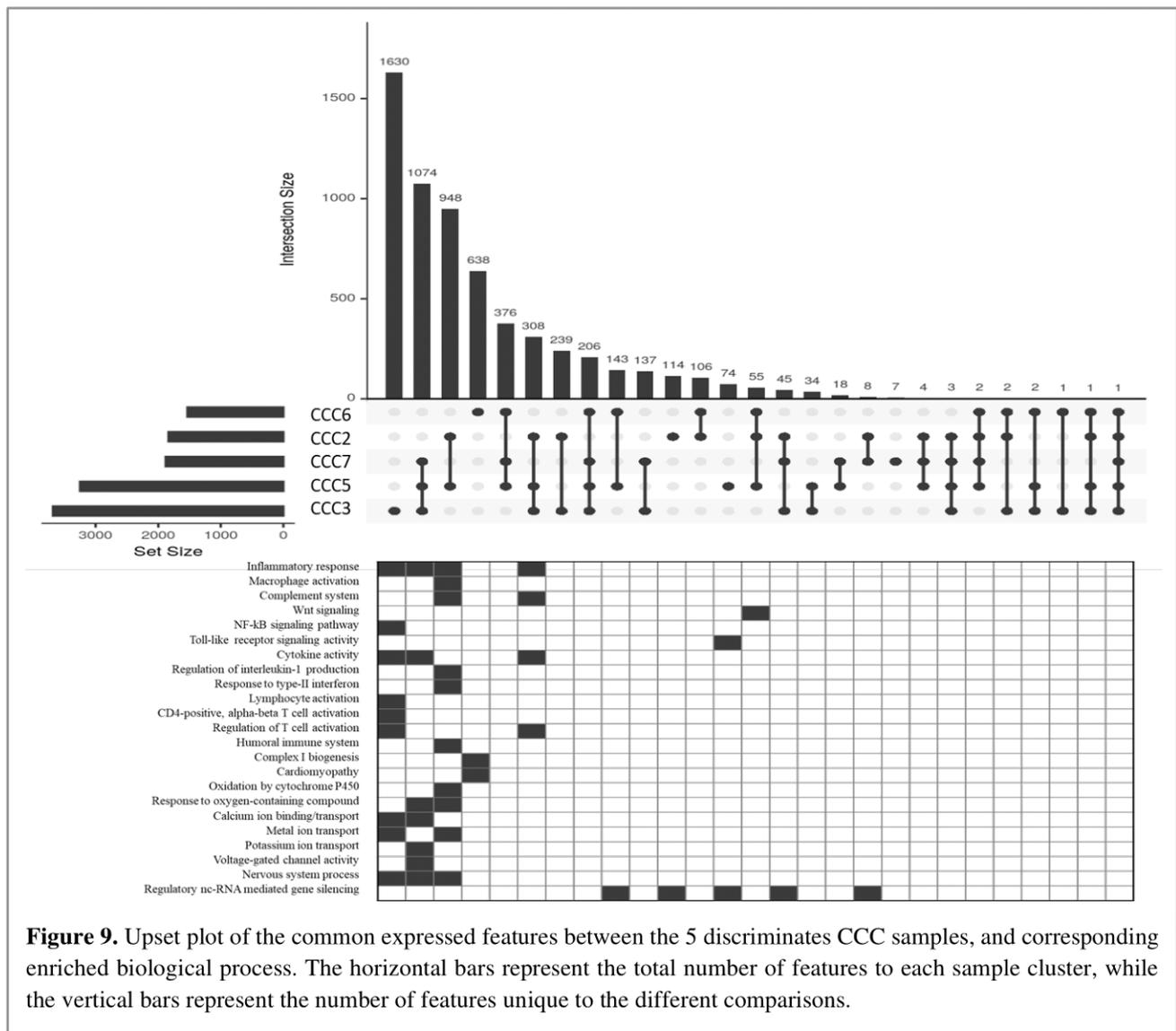


Figure 9. Upset plot of the common expressed features between the 5 discriminates CCC samples, and corresponding enriched biological process. The horizontal bars represent the total number of features to each sample cluster, while the vertical bars represent the number of features unique to the different comparisons.

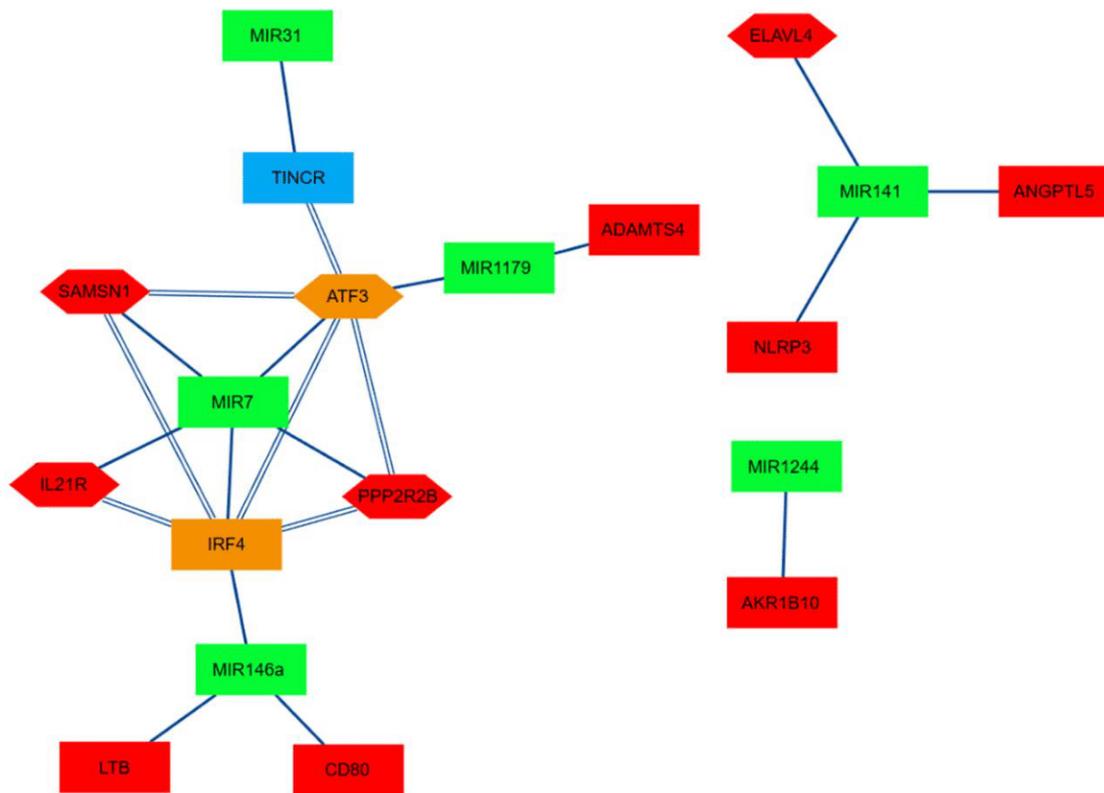


Figure 10. Correlation network of genes only associated to CCC3 cluster. Nodes color: red: protein-coding genes, blue: lncRNA, green: miRNA, orange: transcription factor. Nodes shape: square: feature expression, hexagon: feature expression carrying at least one CpG selected with MOFA+ in its promoter. Edge line color: blue: positively correlated, green: negatively correlated. Edges line type: line: correlated in CCC network, dot: correlated in CTRL network, double line: correlated in both networks.

IRF4 exacerbation as a marker of CCC inflammation

As before, a correlation network was built using the features only associated to CCC3 sample. The resulting network illustrates higher expression on IRF4 transcription factor, as well as some ncRNA including MIR31, TINCR, IRF4, MIR146a, MIR141 and MIR1224. While half of the observed correlations are shared by all samples, CCC3 exhibits a unique miRNA profile, potentially up-regulated genes involved in CCC pathogenesis such as NLRP3 or IL21R (**figure 10**). Interestingly, two miRNAs are correlated to IRF4 and may enhance its expression, resulting in a worse B/T cell inflammatory response.

Both innate and adaptive immune response leads to CCC development

To confirm our previous findings, we conducted a comparative analysis of key features across all CCC samples, with a particular focus on CCC3 and CCC6 (**figure 11**). Notably, CCC3 and CCC6 exhibited elevated levels of both SPIC and TLR7. However, CCC6 displayed a higher expression of TLR8, BLNK, and CD19. In contrast, CCC6 exhibited higher expression levels of TLR10, IFNG, IRF4, NLRP3, and CD8A. Examining the BCR/TCR repertoire, CCC6 stood out as having one of the highest numbers of clonotypes. While this may suggest a greater presence of B/T cells in the CCC myocardium, these results may also indicate a more robust humoral response in CCC6, leading to an elevated expression of IFNG. Furthermore, CCC6 exhibited increased expression of both TLR7 and TLR8, which are endosomal TLRs, along with genes

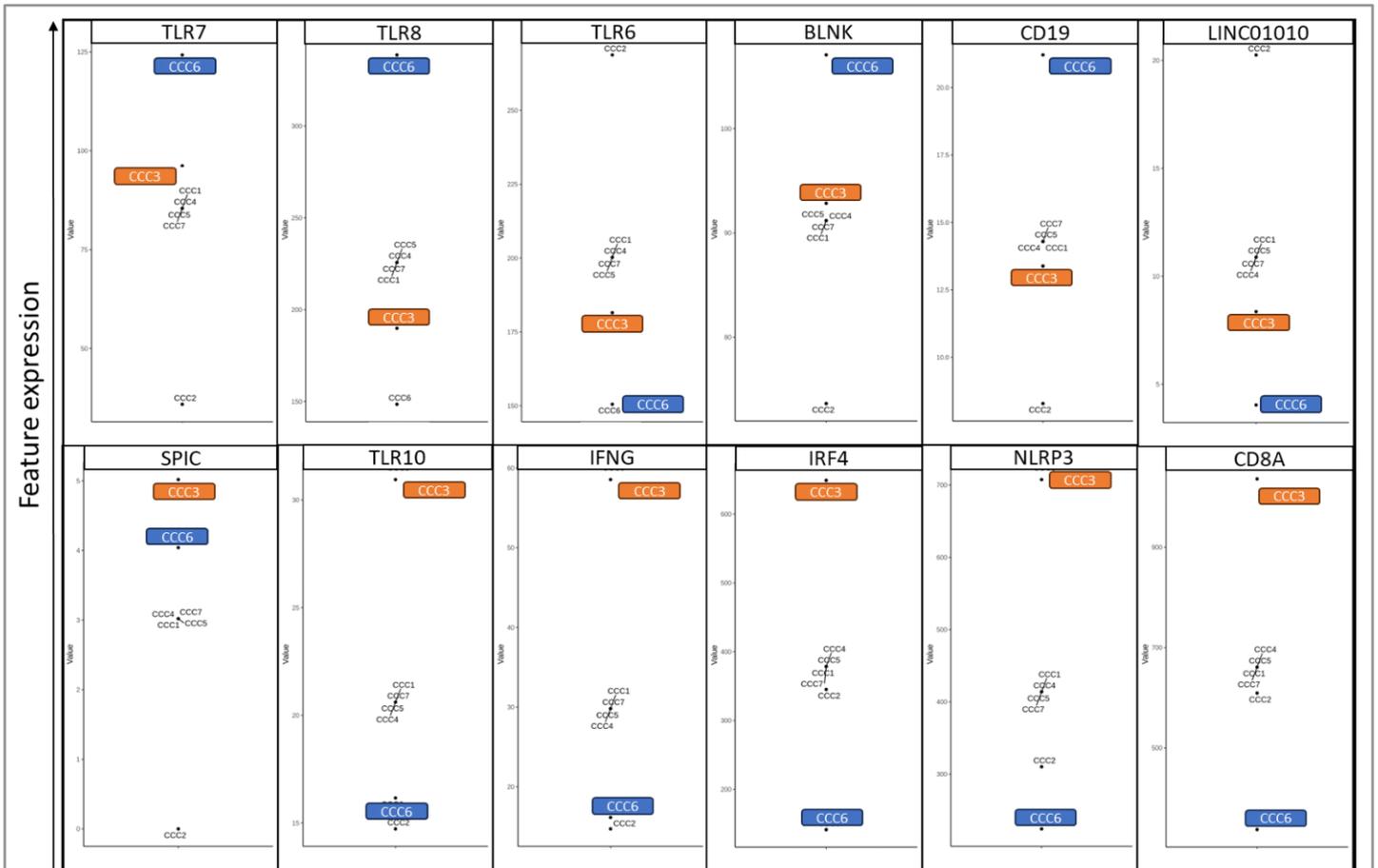


Figure 11. Expression level of the identified features of interest in the previous results, with CCC3 in orange and CCC6 in blue.

related to B cells, underscoring an internal response. While BCR convergence was observed in all samples, both CCC3 and CCC6 resulted in heart tissue damage through distinct processes.

Discussion

In this study, we explored potential factors influencing gene expression in CCC at different omics levels. This multi-omics integration was conducted in a small and incomplete dataset, primarily because of challenges in obtaining human heart samples. To counteract these difficulties, the decision was made to impute the missing values using the entire multi-omics dataset. This approach produced convincing results, with the elimination of samples having minimal impact on the characteristics selected following MOFA analysis. The choice to avoid existing methods for creating synthetic samples, such as SMOTE and KNNOR, was influenced by the complexity of processing different omics data. These methods rely on k-

nearest neighbors to generate synthetic samples, which can be particularly challenging in a multi-omic dataset where each omics has different number of features. Our approach has overcome these difficulties by avoiding the need for such neighbor selection. In single-omic datasets, the results are very close to those obtained with SMOTE, which reinforces the preference for this solution.

Based on genomic, transcriptomic and methylomic dataset, the multi-omics integration has highlighted 1,214 genes, 2,066 methylation site and 58 miRNA discriminating healthy controls to CCC patients. Moreover, 488/1272 expressed features (46%) were never associated to this disease, according to ChagasDB (9). The genes were mostly associated to immune-related processes, mostly involved in inflammatory response and Th1 cell activation. The involvement of Th1-cell related genes in CCC has already been largely demonstrated in previous study (10,11,14).

In order to understand the relations between these features, two correlation networks were built, each specific to one phenotype. The same features exhibit contrasting behaviors between healthy control and CCC patient networks. In CCC, three regulatory clusters were observed, including genes involved in T-cell/TCR processes, or BCR and TLR pathways. We have identified several transcription factors correlated with genes involved in opposite biological processes, notably pro- and anti-inflammatory response. Among them, IRF4 may act with different TF, resulting in different functions. IRF4 is a transcription factor mostly produced by B cell after BCR stimulation by antigens and co-stimulatory signal with T-cell (37) but also in T cell after TCR stimulation. In our data, IRF4 is correlated with both TCR and BCR genes. In vitro, IRF4 overexpression leads to inflammasome inhibition (38), while NLRP3 is overexpressed in CCC samples, suggesting a default in inflammatory regulation. Interestingly, its impact on the immune response depends on its interactors and the cell type in which it is produced (39). Notably, in T-cell, it interacts with BATF, and in B cell, with SPIB. In Treg specially, it has been shown to interact with FOXP3. In our data, IRF4 is correlated with FOXP3 in CTRL network only, and with BATF (AP-1 family) in CCC network. Recently, Arnold PR. et al demonstrate that BATF3, a TF of AP-1 family, partnering with IRF4 to repress FOXP3, and so Treg production (40). On the opposite, both IRF4 and BRDM1 are correlated with CTLA4, a Treg marker, and with SOCS3, a regulator of STAT3 and downregulated in CCC. The reduction of SOCS3 gene expression was associated to IL-10 expression and Th2 profile induction (41). SOCS3 is also negative correlated with miR-551b, an activator of STAT3 (42). In our data, STAT3 is under-expressed in CCC compared to CTRL, possibly resulting from a lack of activation by miR-551b. Furthermore, various studies have linked IRF4 to numerous transcription factors observed in CCC, such as RUNX3 (43), PAX5 (44), IKZF1 (45), TBX21, and EOMES (46). Our analysis suggests that RUNX3 plays a pivotal role in activating the inflammatory response, potentially influenced by DNA methylation within its target promoter region. A prior study from our research predicted that both PAX5 and IRF4 may bind to the RUNX3 promoter,

with their binding potentially regulated by differences in methylation levels (11). Additionally, our in vitro experiments demonstrated that methylation of the RUNX3 promoter impacts its expression (14). Within the CCC network, we observed that CCC6 sample is characterized by elevated IRF4 expression, along with increased levels of IFN- γ and the NLRP3 inflammasome. Considering the CD8A and CD4 gene markers, as well as a high number of TCR clonotypes, it is plausible that IRF4 is associated with an intensified T-cell response, leading to cardiac damage. In this line, IRF4 exhibits correlations with LINC00964 and LINC00528, both of which have already been implicated in heart diseases (47,48). MIAT, a lncRNA previously linked to CCC (13) and present in our network, is also overexpressed in CCC6 patient. Collectively, these findings strongly indicate that IRF4 serves as a key regulator of the inflammatory response.

Our analysis has also revealed some TLR overexpressed in CCC patients, including TLR6, TLR7, TLR8 and TLR10. TLR10 is an endosomal TLR mainly located in B cell (49), potentially activated by double-stranded DNA. Although little known, TLR10 seems to differ from other TLRs in its anti-inflammatory action (50). However, in our data, TLR10 is correlated with pro-inflammatory genes, TLR4 and BCR-related genes (CD79A, BANK1), suggesting a dual function of this TLR. On the other hand, TLR7 is correlated with SPIC, and both are linked to immune regulation genes. In vitro and in vivo experiments highlights the SPIC activation through the TLR/NF- κ B pathway (51), inducing the production of anti-inflammatory cytokines. In this line, SPIC is also correlated to IL-4 and IL-411, a cytokine involved in Th2 differentiation (52), as well as FRZB, a regulator of Wnt pathway (53,54). We can hypothesize here that SPIC may play an inflammatory regulation process, while not sufficient to reduce the inflammation in CCC. It is interesting to note that both TLR7 and SPIC are up regulated in CCC3 and CCC6 patients. On the other hand, TLR8 is positively correlated with inflammatory genes, and up regulated in CCC6 sample. TLR8 is an endosomal TLR activated by single-stranded DNA. Considering that *T.cruzi* DNA is single-stranded, it may be directly activated by

resting amastigotes in cardiomyocytes. However, CCC6 samples present mitochondrial dysfunctions, potentially due to several mitochondria mutations. Mitochondrial dysfunctions are known to participate to CCC pathogenic processes. The accumulation of defective mitochondria may lead to mitophagy (55), resulting in the generation of mitochondria DAMP, which can activate endosomal TLR (56). While mtDNA is double-stranded, RNA:DNA hybrids form produced during the transcription could be recognized by TLR (57).

Regarding specifically at CCC6 mtDNA genotype, the rs28647976 mutation, a missense variant in MT-ND1, a gene involved in complex I, was associated to miR-369-5p and miR-591. The overexpression of miR-369 conduct to a reduction of inflammation in rat (58). Moreover, miR-369-5p target DNMT3 gene, conducting in reduction of cardiac fibroblasts proliferation (59). The miR-369-3p is associated to a reduction of NO production (60) and inflammasome pathway (61). In CCC context, may be attempting to mitigate the oxidative stress that is exacerbated by both CCC pathogenic process and MT-ND1 mutation, and therefore be an interesting therapeutic target. The identification of specific mutations on some samples are in line with the family-specific mutations previously identified (62).

While different pathogenic pathways have been identified in different CCC patients, the BCR repertoire analysis support the existence of specific selection of BCR light chains in the absence of the pathogen in all CCC samples. The convergence towards sequences common to all the patients studied indicates the possibility of the establishment of B cell autoreactivity against similar host proteins, in line with previous observations in humans and mice (8). Interestingly, IRF4 is required to perform the light chain rearrangements (63). In addition, some study reveals the involvement of both BCR and TLR in the development of auto-immune disease (64,65). Our multi-omics integration identified several features known to be involved in autoimmune disease, including FCRL5, one of the top discriminative features (66). The hypothesis of an autoimmune reaction occurring in CCC heart tissue has already been supported by different studies. For example, T-cell obtains from CCC patients reacts after stimulation by T.cruzi derived

peptides, like the B13 protein (67). Considering the absence of TCR public repertoire in our sample, the specificity of TCR, and the resulting T-cell response, may be variable on the different samples. This result could also be due to an insufficient sequencing depth, and BCR/TCR specific analysis should be performed.

To conclude, this analysis has identified numerous regulatory elements probably involved in CCC pathogenesis. While all CCC samples present a similar B-cell response, some of them develop specific process, leading to the heart cardiomyopathy. Considering the opposite roles mediated by TF in different cell types, further analysis at single-cell levels are required to conclude about the importance of anti-inflammatory processes observed in CCC, and their lack of efficiency. Nonetheless, the observations conducted in this study could help, in the future, to develop effective treatments depending on the divergent reactions of each patient.

Funding

This work was supported by the Institut National de la Santé et de la Recherche Médicale (INSERM); the Aix-Marseille University (grant number: AMIDEX “International_2018” MITOMUTCHAGAS); the French Agency for Research (Agence Nationale de la Recherche-ANR (grant numbers: “Br-Fr-Chagas”, “landscardio”); the CNPq (Brazilian Council for Scientific and Technological Development); and the FAPESP (São Paulo State Research Funding Agency Brazil (grant numbers: 2013/50302-3, 2014/50890-5); the National Institutes of Health/USA (grant numbers: 2 P50 AI098461-02 and 2U19AI098461-06). This work was founded by the Inserm Cross-Cutting Project GOLD. This project has received funding from the Excellence Initiative of Aix-Marseille University - A*Midex a French “Investissements d’Avenir programme”- Institute MarMaRa AMX-19-IET-007. JPSN was a recipient of a MarMaRa fellowship. EC-N and JK are recipients of productivity awards by CNPq. The funders did not play any role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgments

Center de Calcul Intensif d'Aix-Marseille is acknowledged for granting access to its high-performance computing resources.

References

1. Irish A, Whitman JD, Clark EH, Marcus R, Bern C. Updated Estimates and Mapping for Prevalence of Chagas Disease among Adults, United States. *Emerg Infect Dis.* juill 2022;28(7):1313-20.
2. Antinori S, Galimberti L, Bianco R, Grande R, Galli M, Corbellino M. Chagas disease in Europe: A review for the internist in the globalized world. *European Journal of Internal Medicine.* 1 sept 2017;43:6-15.
3. Iglesias Rodríguez IM, Miura S, Maeda T, Imai K, Smith C, Vasquez Velasquez C, et al. Analysis of the Chagas disease situation in Japan: a cross sectional study and cost-effectiveness analysis of a Chagas disease screening program. *Lancet Reg Health West Pac.* 5 sept 2022;31:100574.
4. Pérez-Molina JA, Crespillo-Andújar C, Bosch-Nicolau P, Molina I. Trypanocidal treatment of Chagas disease. *Enfermedades infecciosas y microbiología clínica (English ed).* 1 nov 2021;39(9):458-70.
5. Nunes MCP, Beaton A, Acquatella H, Bern C, Bolger AF, Echeverría LE, et al. Chagas Cardiomyopathy: An Update of Current Clinical Knowledge and Management: A Scientific Statement From the American Heart Association. *Circulation.* 18 sept 2018;138(12):e169-209.
6. Cristovão-Silva AC, Brelaz-de-Castro MCA, Hernandez MZ, Pereira VRA. Chagas disease: Immunology of the disease at a glance. *Cytokine & Growth Factor Reviews.* 1 déc 2021;62:15-22.
7. Wen J jun, Yachelini PC, Sembaj A, Manzur RE, Garg NJ. Increased oxidative stress is correlated with mitochondrial dysfunction in chagasic patients. *Free Radical Biology and Medicine.* 15 juill 2006;41(2):270-6.
8. De Bona E, Lidani KCF, Bavia L, Omidian Z, Gremski LH, Sandri TL, et al. Autoimmunity in Chronic Chagas Disease: A Road of Multiple Pathways to Cardiomyopathy? *Front Immunol.* 6 août 2018;9:1842.
9. Brochet P, Mouren JC, Hannouche L, Lopez F, Ballester B, Cunha-Neto E, et al. ChagasDB: 80 years of publicly available data on the molecular host response to *Trypanosoma cruzi* infection in a single database. *Database (Oxford).* 26 mai 2023;2023:baad037.
10. Cunha-Neto E, Dzau VJ, Allen PD, Stamatiou D, Benvenuti L, Higuchi ML, et al. Cardiac Gene Expression Profiling Provides Evidence for Cytokinopathy as a Molecular Mechanism in Chagas' Disease Cardiomyopathy. *Am J Pathol.* août 2005;167(2):305-13.
11. Brochet P, Ianni BM, Laugier L, Frade AF, Silva Nunes JP, Teixeira PC, et al. Epigenetic regulation of transcription factor binding motifs promotes Th1 response in Chagas disease cardiomyopathy. *Front Immunol.* 22 août 2022;13:958200.
12. Laugier L, Ferreira LRP, Ferreira FM, Cabantous S, Frade AF, Nunes JP, et al. miRNAs may play a major role in the control of gene expression in key pathobiological processes in Chagas disease cardiomyopathy. *PLoS Negl Trop Dis.* 22 déc 2020;14(12):e0008889.
13. Frade AF, Laugier L, Ferreira LRP, Baron MA, Benvenuti LA, Teixeira PC, et al. Myocardial Infarction-Associated Transcript, a Long Noncoding RNA, Is Overexpressed During Dilated Cardiomyopathy Due to Chronic Chagas Disease. *The Journal of Infectious Diseases.* 1 juill 2016;214(1):161-5.
14. Laugier L, Frade AF, Ferreira FM, Baron MA, Teixeira PC, Cabantous S, et al. Whole-Genome Cardiac DNA Methylation Fingerprint and Gene Expression Analysis Provide New Insights in the Pathogenesis of Chronic Chagas Disease Cardiomyopathy. *Clin Infect Dis.* 1 oct 2017;65(7):1103-11.
15. Chen G, Wang Z, Wang D, Qiu C, Liu M, Chen X, et al. LncRNADisease: a database for long-non-coding RNA-associated diseases. *Nucleic Acids Res.* janv 2013;41(Database issue):D983-6.
16. Argelaguet R, Arnol D, Bredikhin D, Deloro Y, Velten B, Marioni JC, et al. MOFA+: a statistical framework for comprehensive integration of multi-modal single-cell data. *Genome Biol.* 11 mai 2020;21:111.
17. Du P, Zhang X, Huang CC, Jafari N, Kibbe WA, Hou L, et al. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics.* 30 nov 2010;11(1):587.
18. Gillespie M, Jassal B, Stephan R, Milacic M, Rothfels K, Senff-Ribeiro A, et al. The reactome pathway knowledgebase 2022. *Nucleic Acids Res.* 12 nov 2021;50(D1):D687-92.
19. Weinstein JN, Collisson EA, Mills GB, Shaw KM, Ozenberger BA, Ellrott K, et al. The Cancer Genome Atlas Pan-Cancer Analysis Project. *Nat Genet.* oct 2013;45(10):1113-20.

20. Blagus R, Lusa L. SMOTE for high-dimensional class-imbalanced data. *BMC Bioinformatics*. 22 mars 2013;14:106.
21. Islam A, Belhaouari SB, Rehman AU, Bensmail H. KNNOR: An oversampling technique for imbalanced datasets. *Applied Soft Computing*. 1 janv 2022;115:108288.
22. Singh A, Shannon CP, Gautier B, Rohart F, Vacher M, Tebbutt SJ, et al. DIABLO: an integrative approach for identifying key molecular drivers from multi-omics assays. *Bioinformatics*. 1 sept 2019;35(17):3055-62.
23. Lock EF, Hoadley KA, Marron JS, Nobel AB. JOINT AND INDIVIDUAL VARIATION EXPLAINED (JIVE) FOR INTEGRATED ANALYSIS OF MULTIPLE DATA TYPES. *Ann Appl Stat*. 1 mars 2013;7(1):523-42.
24. Voillet V, Besse P, Liaubet L, San Cristobal M, González I. Handling missing rows in multi-omics data integration: multiple imputation in multiple factor analysis framework. *BMC Bioinformatics*. 3 oct 2016;17:402.
25. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res*. 8 janv 2019;47(Database issue):D607-13.
26. Oughtred R, Rust J, Chang C, Breitkreutz B, Stark C, Willems A, et al. The BioGRID database: A comprehensive biomedical resource of curated protein, genetic, and chemical interactions. *Protein Sci*. janv 2021;30(1):187-200.
27. Yang S, Kim CY, Hwang S, Kim E, Kim H, Shim H, et al. COEXPEDIA: exploring biomedical hypotheses via co-expressions associated with medical subject headings (MeSH). *Nucleic Acids Res*. 4 janv 2017;45(Database issue):D389-96.
28. Fang L, Li Y, Ma L, Xu Q, Tan F, Chen G. GRNdb: decoding the gene regulatory networks in diverse human and mouse conditions. *Nucleic Acids Res*. 5 nov 2020;49(D1):D97-103.
29. del Toro N, Shrivastava A, Ragueneau E, Meldal B, Combe C, Barrera E, et al. The IntAct database: efficient access to fine-grained molecular interaction data. *Nucleic Acids Res*. 11 nov 2021;50(D1):D648-53.
30. Zhao H, Yin X, Xu H, Liu K, Liu W, Wang L, et al. LncTarD 2.0: an updated comprehensive database for experimentally-supported functional lncRNA–target regulations in human diseases. *Nucleic Acids Res*. 2 nov 2022;51(D1):D199-207.
31. Xiao F, Zuo Z, Cai G, Kang S, Gao X, Li T. miRecords: an integrated resource for microRNA–target interactions. *Nucleic Acids Res*. janv 2009;37(Database issue):D105-10.
32. Huang HY, Lin YCD, Cui S, Huang Y, Tang Y, Xu J, et al. miRTarBase update 2022: an informative resource for experimentally validated miRNA–target interactions. *Nucleic Acids Res*. 30 nov 2021;50(D1):D222-30.
33. Karagkouni D, Paraskevopoulou MD, Chatzopoulos S, Vlachos IS, Tastsoglou S, Kanellos I, et al. DIANA-TarBase v8: a decade-long collection of experimentally supported miRNA–gene interactions. *Nucleic Acids Res*. 4 janv 2018;46(Database issue):D239-45.
34. Song L, Cohen D, Ouyang Z, Cao Y, Hu X, Liu XS. TRUST4: immune repertoire reconstruction from bulk and single-cell RNA-seq data. *Nat Methods*. juin 2021;18(6):627-30.
35. Rich-Griffin C. immunarch - An R Package for Painless Bioinformatics Analysis of T-cell and B-cell Immune Repertoire Data [Internet]. 2023 [cité 15 oct 2023]. Disponible sur: https://github.com/crichgriffin/immunarch_code
36. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol*. 11 oct 2011;7:539.
37. Shukla V, Lu R. IRF4 and IRF8: Governing the virtues of B Lymphocytes. *Front Biol (Beijing)*. août 2014;9(4):269-82.
38. McDaniel MM, Kottyan LC, Singh H, Pasare C. Suppression of Inflammasome Activation by IRF8 and IRF4 in cDCs Is Critical for T Cell Priming. *Cell Rep*. 5 mai 2020;31(5):107604.
39. Huber M, Lohoff M. IRF4 at the crossroads of effector T-cell fate decision. *Eur J Immunol*. juill 2014;44(7):1886-95.
40. Arnold PR, Wen M, Zhang L, Ying Y, Xiao X, Chu X, et al. Suppression of FOXP3 expression by the AP-1 family transcription factor BATF3 requires partnering with IRF4. *Front Immunol*. 25 août 2022;13:966364.
41. Kinjyo I, Inoue H, Hamano S, Fukuyama S, Yoshimura T, Koga K, et al. Loss of SOCS3 in T helper cells resulted in reduced immune responses and hyperproduction of interleukin 10 and transforming growth factor– β 1. *J Exp Med*. 17 avr 2006;203(4):1021-31.
42. Chaluvally-Raghavan P, Jeong KJ, Pradeep S, Silva AM, Yu S, Liu W, et al. Direct upregulation of STAT3 by microRNA-551b-3p deregulates growth and metastasis of ovarian cancer. *Cell Rep*. 17 mai 2016;15(7):1493-504.
43. Cao Y, Li H, Sun Y, Chen X, Liu H, Gao X, et al. Interferon regulatory factor 4 regulates

thymocyte differentiation by repressing Runx3 expression. *European Journal of Immunology*. 2010;40(11):3198-209.

44. Teng Y, Takahashi Y, Yamada M, Kurosu T, Koyama T, Miura O, et al. IRF4 negatively regulates proliferation of germinal center B cell-derived Burkitt's lymphoma cell lines and induces differentiation toward plasma cells. *European Journal of Cell Biology*. 19 oct 2007;86(10):581-9.

45. Ochiai K, Kondo H, Okamura Y, Shima H, Kurokuchi Y, Kimura K, et al. Zinc finger-IRF composite elements bound by Ikaros/IRF4 complexes function as gene repression in plasma cell. *Blood Adv*. 18 avr 2018;2(8):883-94.

46. Nayar R, Schutten E, Jangalwe S, Durost PA, Kenney LL, Conley JM, et al. IRF4 Regulates the Ratio of T-Bet to Eomesodermin in CD8+ T Cells Responding to Persistent LCMV Infection. *PLoS One*. 29 déc 2015;10(12):e0144826.

47. Liu Y, Liu N, Bai F, Liu Q. Identifying ceRNA Networks Associated With the Susceptibility and Persistence of Atrial Fibrillation Through Weighted Gene Co-Expression Network Analysis. *Front Genet*. 23 juin 2021;12:653474.

48. Liu K, Zhao D, Wang D. LINC00528 regulates myocardial infarction by targeting the miR-143-3p/COX-2 axis. *Bioengineered*. 27 déc 2019;11(1):11-8.

49. Lee SMY, Yip TF, Yan S, Jin DY, Wei HL, Guo RT, et al. Recognition of Double-Stranded RNA and Regulation of Interferon Pathway by Toll-Like Receptor 10. *Front Immunol*. 16 mars 2018;9:516.

50. Fore F, Indriputri C, Mamutse J, Nugraha J. TLR10 and Its Unique Anti-Inflammatory Properties and Potential Use as a Target in Therapeutics. *Immune Netw*. 23 avr 2020;20(3):e21.

51. Alam Z, Devalaraja S, Li M, To TKJ, Folkert IW, Mitchell-Velasquez E, et al. Counter Regulation of Spic by NF- κ B and STAT Signaling Controls Inflammation and Iron Metabolism in Macrophages. *Cell Rep*. 30 juin 2020;31(13):107825.

52. Chen L, Grabowski KA, Xin JP, Coleman J, Huang Z, Espiritu B, et al. IL-4 induces differentiation and expansion of Th2 cytokine-producing eosinophils. *J Immunol*. 15 févr 2004;172(4):2059-66.

53. Leijten JC, Bos SD, Landman EB, Georgi N, Jahr H, Meulenbelt I, et al. GREM1, FRZB and DKK1 mRNA levels correlate with osteoarthritis and are regulated by osteoarthritis-associated factors. *Arthritis Res Ther*. 2013;15(5):R126.

54. Huang X, Zhong L, van Helvoort E, Lafeber F, Mastbergen S, Hendriks J, et al. The Expressions

of Dickkopf-Related Protein 1 and Frizzled-Related Protein Are Negatively Correlated to Local Inflammation and Osteoarthritis Severity. *Cartilage*. oct 2021;12(4):496-504.

55. Andrieux P, Chevillard C, Cunha-Neto E, Nunes JPS. Mitochondria as a Cellular Hub in Infection and Inflammation. *Int J Mol Sci*. 20 oct 2021;22(21):11338.

56. Grazioli S, Pugin J. Mitochondrial Damage-Associated Molecular Patterns: From Inflammatory Signaling to Human Diseases. *Front Immunol*. 4 mai 2018;9:832.

57. Riley JS, Tait SW. Mitochondrial DNA in inflammation and immunity. *EMBO Rep*. 3 avr 2020;21(4):e49799.

58. Wang J, Chen X, Huang W. MicroRNA-369 attenuates hypoxia-induced cardiomyocyte apoptosis and inflammation via targeting TRPV3. *Braz J Med Biol Res*. 15 janv 2021;54(3):e10550.

59. Tao H, Dai C, Ding JF, Yang JJ, Ding XS, Xu SS, et al. Epigenetic aberrations of miR-369-5p and DNMT3A control Patched1 signal pathway in cardiac fibrosis. *Toxicology*. 1 déc 2018;410:182-92.

60. Scalavino V, Liso M, Cavalcanti E, Gigante I, Lippolis A, Mastronardi M, et al. miR-369-3p modulates inducible nitric oxide synthase and is involved in regulation of chronic inflammatory response. *Sci Rep*. 29 sept 2020;10(1):15942.

61. Scalavino V, Piccinno E, Valentini AM, Schena N, Armentano R, Giannelli G, et al. miR-369-3p Modulates Intestinal Inflammatory Response via BRCC3/NLRP3 Inflammasome Axis. *Cells*. 31 août 2023;12(17):2184.

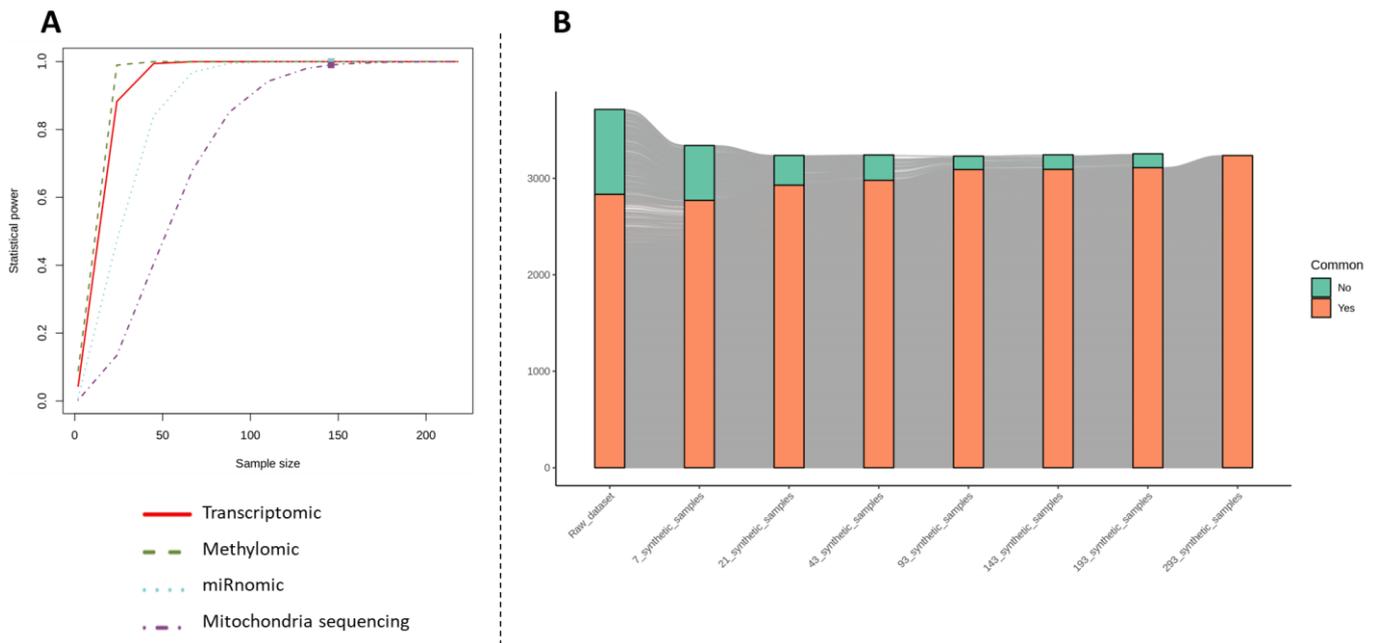
62. Ouarhache M, Marquet S, Frade AF, Ferreira AM, Ianni B, Almeida RR, et al. Rare Pathogenic Variants in Mitochondrial and Inflammation-Associated Genes May Lead to Inflammatory Cardiomyopathy in Chagas Disease. *J Clin Immunol*. 2021;41(5):1048-63.

63. Lu R, Medina KL, Lancki DW, Singh H. IRF-4,8 orchestrate the pre-B-to-B transition in lymphocyte development. *Genes Dev*. 15 juill 2003;17(14):1703-8.

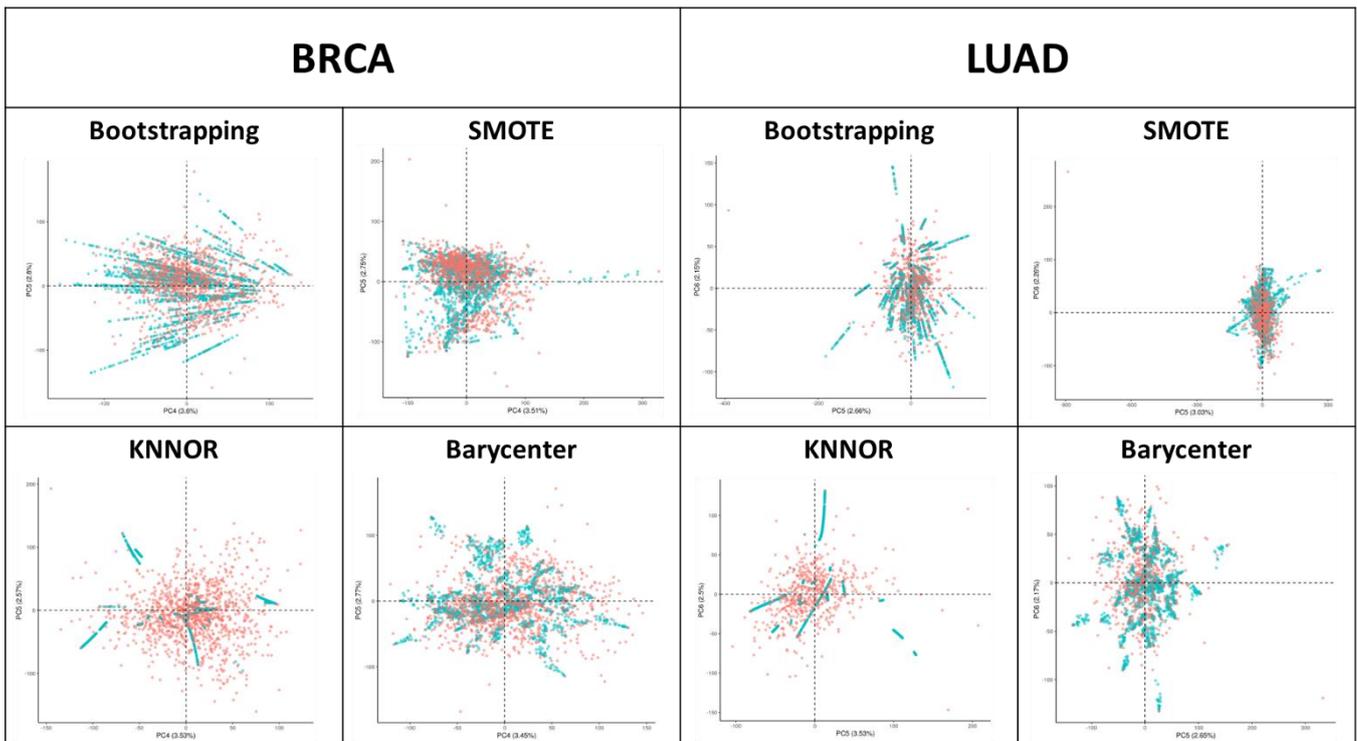
64. Rip J, de Bruijn MJW, Appelman MK, Pal Singh S, Hendriks RW, Corneth OBJ. Toll-Like Receptor Signaling Drives Btk-Mediated Autoimmune Disease. *Front Immunol*. 30 janv 2019;10:95.

65. Weindel CG, Richey LJ, Bolland S, Mehta AJ, Kearney JF, Huber BT. B cell autophagy mediates TLR7-dependent autoimmunity and inflammation. *Autophagy*. 29 juin 2015;11(7):1010-24.

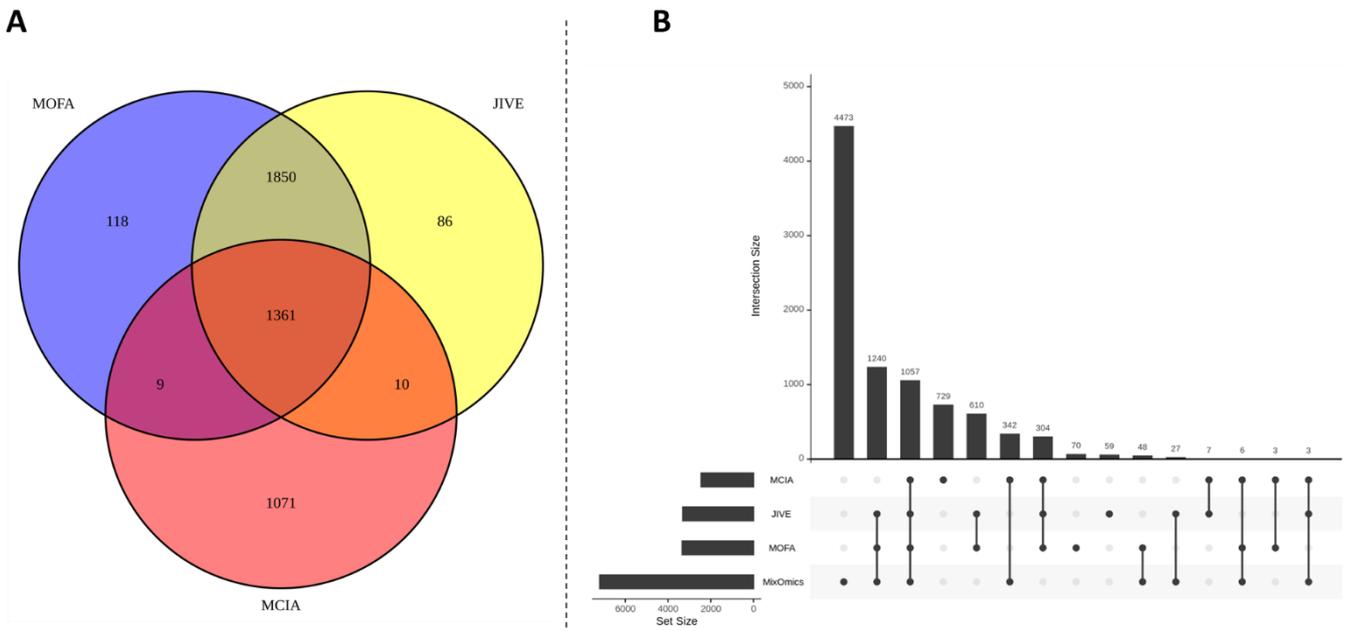
66. Ono C, Tanaka S, Myouzen K, Iwasaki T, Ueda M, Oda Y, et al. Upregulated Fcrl5 disrupts B cell anergy causes autoimmune disease. *Frontiers in Immunology* [Internet]. 2023 [cité 15 oct 2023];14. Disponible sur: <https://www.frontiersin.org/articles/10.3389/fimmu.2023.1276014>
67. Cunha-Neto E, Coelho V, Guilherme L, Fiorelli A, Stolf N, Kalil J. Autoimmunity in Chagas' disease. Identification of cardiac myosin-B13 *Trypanosoma cruzi* protein crossreactive T cell clones in heart lesions of a chronic Chagas' cardiomyopathy patient. *J Clin Invest*. 15 oct 1996;98(8):1709-12.



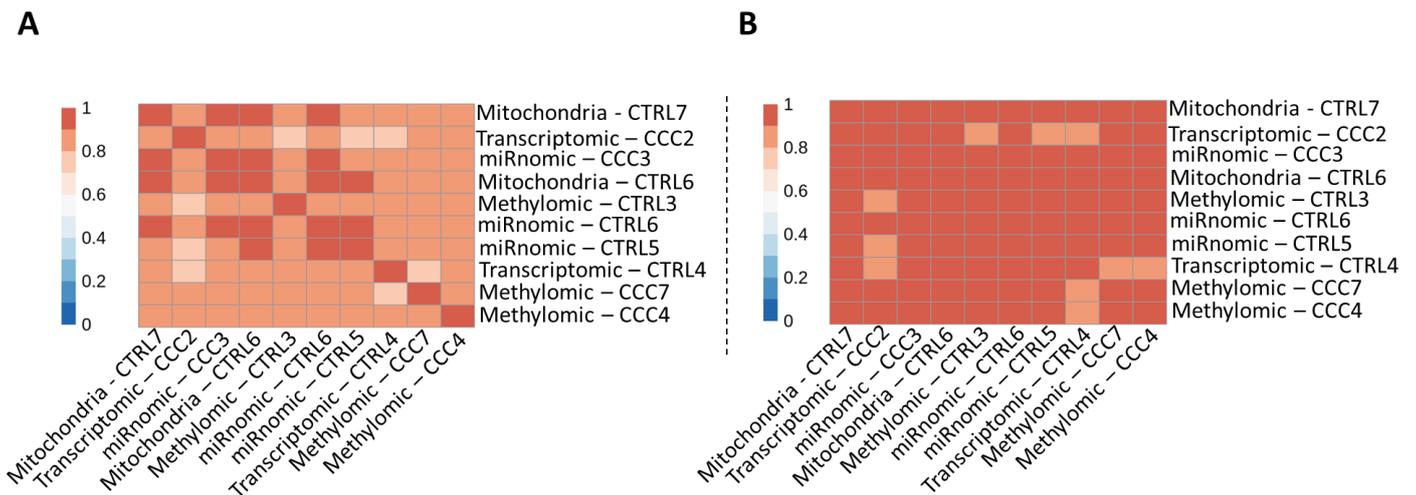
Supplementary Figure 1. (A) Statistical power estimation with the increasing number of samples per omic. **(B)** Alluvial plot of the common selected features with the increasing number of samples.



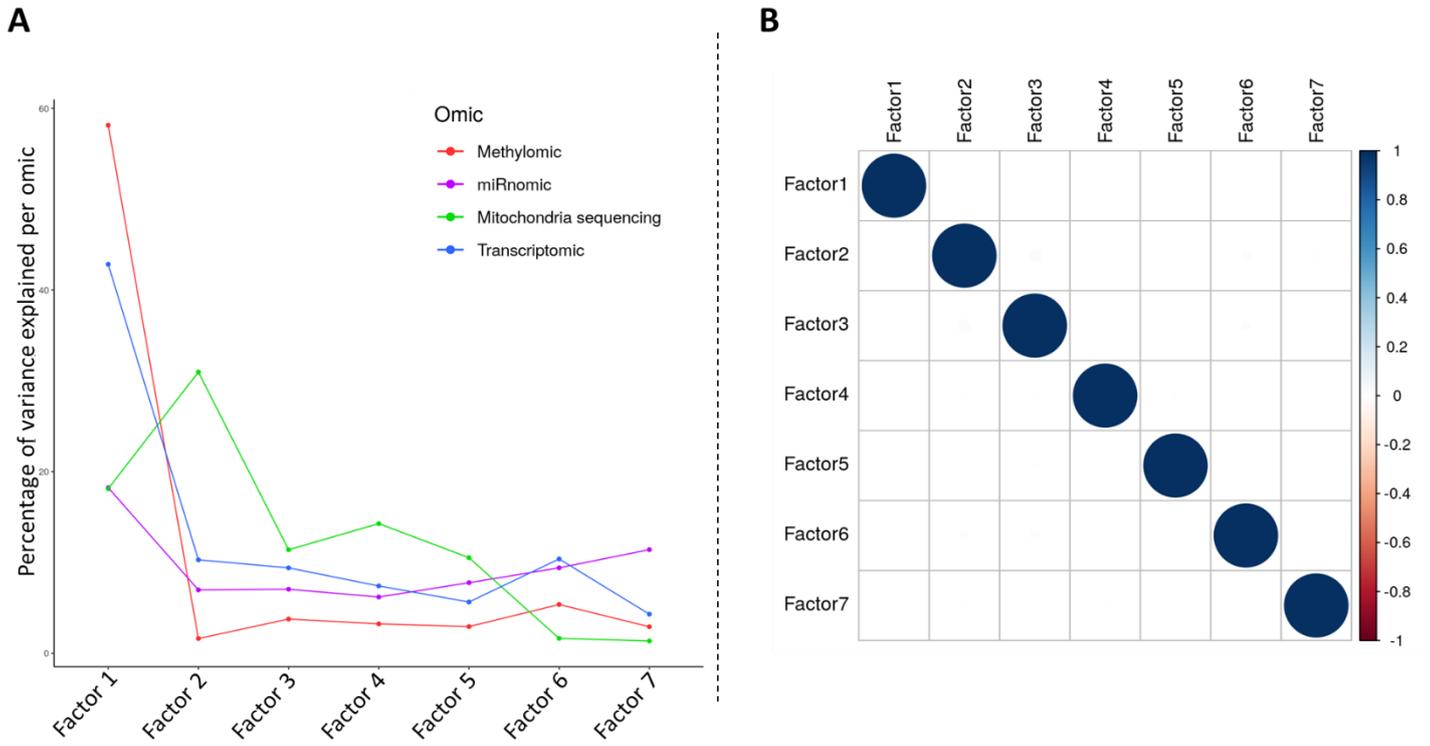
Supplementary Figure 2. Visualization of the dimension 5 and 6 of PCA analysis performed, for each phenotype (BRCA and LUAD) on raw and synthetic samples generated with bootstrapping, SMOTE, KNNOR or barycenter methods (red : raw samples, blue : synthetic samples).



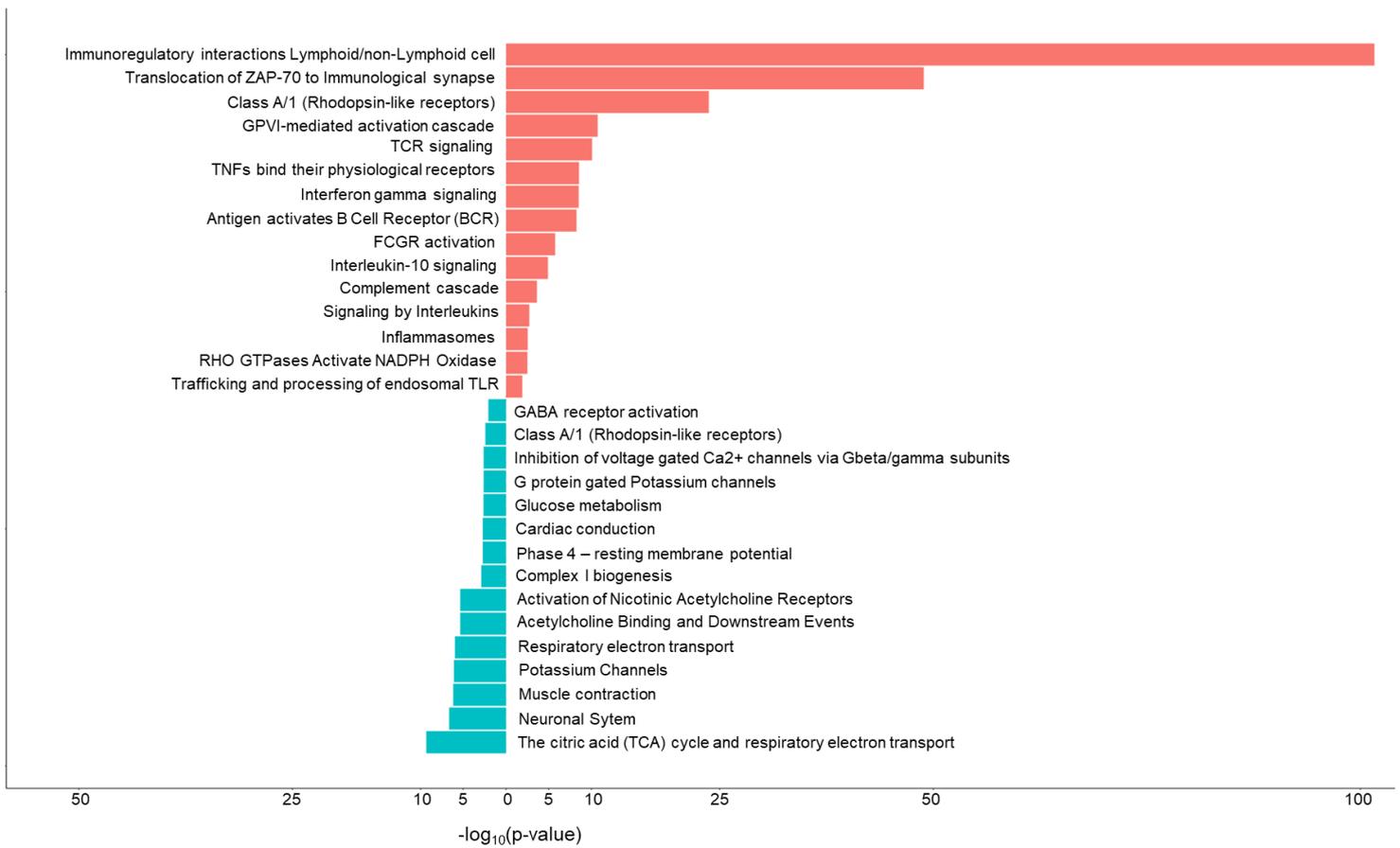
Supplementary Figure 3. Comparison of the selected features after different multi-omics integration tools. **(A)** Venn diagram between MOFA+, MCIA and JIVE. **(B)** Upset plot comparing the features in common between MOFA+, MCIA, JIVE and MixOmics integration. The horizontal bars represent the total number of features per method, while the vertical bars represent the number of features unique to the different comparisons.



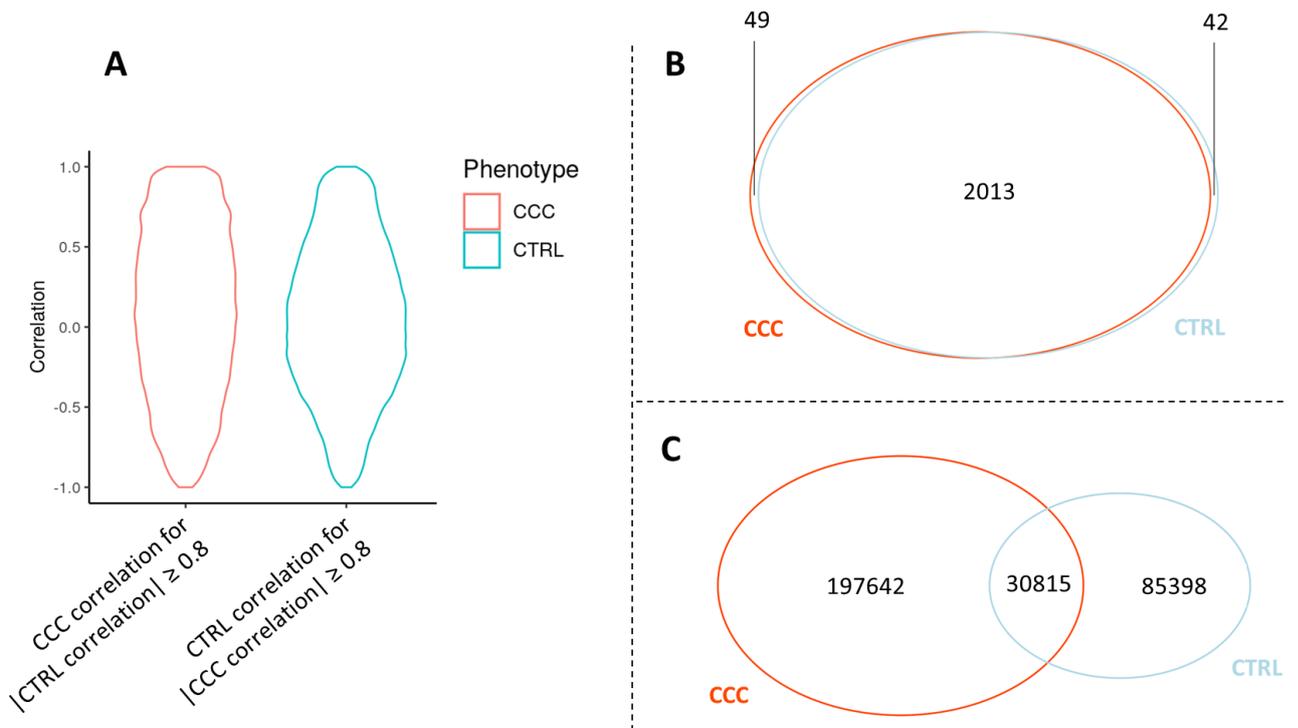
Supplementary Figure 4. Comparison of the selected features after different iteration, removing each type one raw sample in one omic. **(A)** Jaccard index, **(B)** minimal overlap.



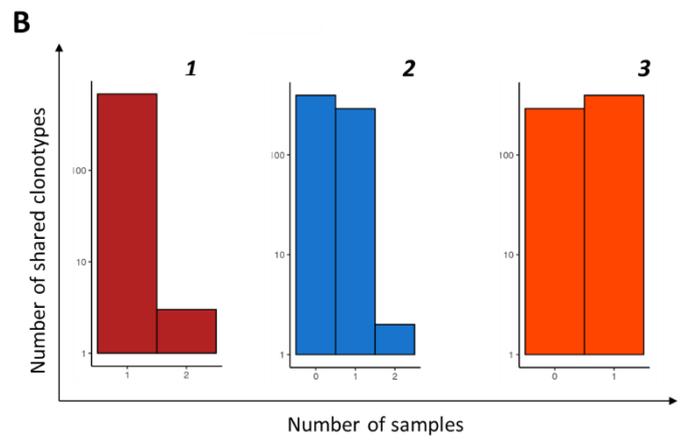
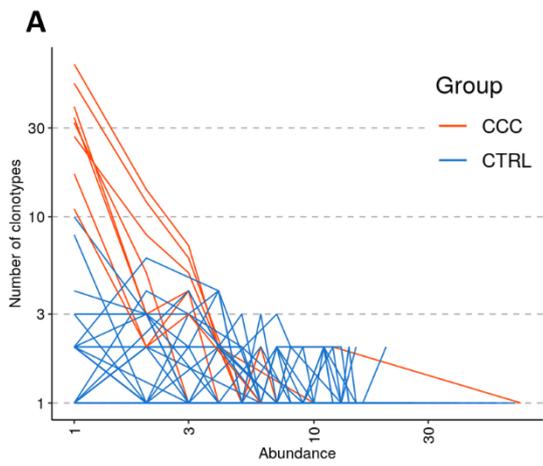
Supplementary Figure 5. Evaluation of the 7-factor selected with CTRL and CCC samples. **(A)** Percentage of variance explained per omic and per factor. **(B)** Correlation between each factor.



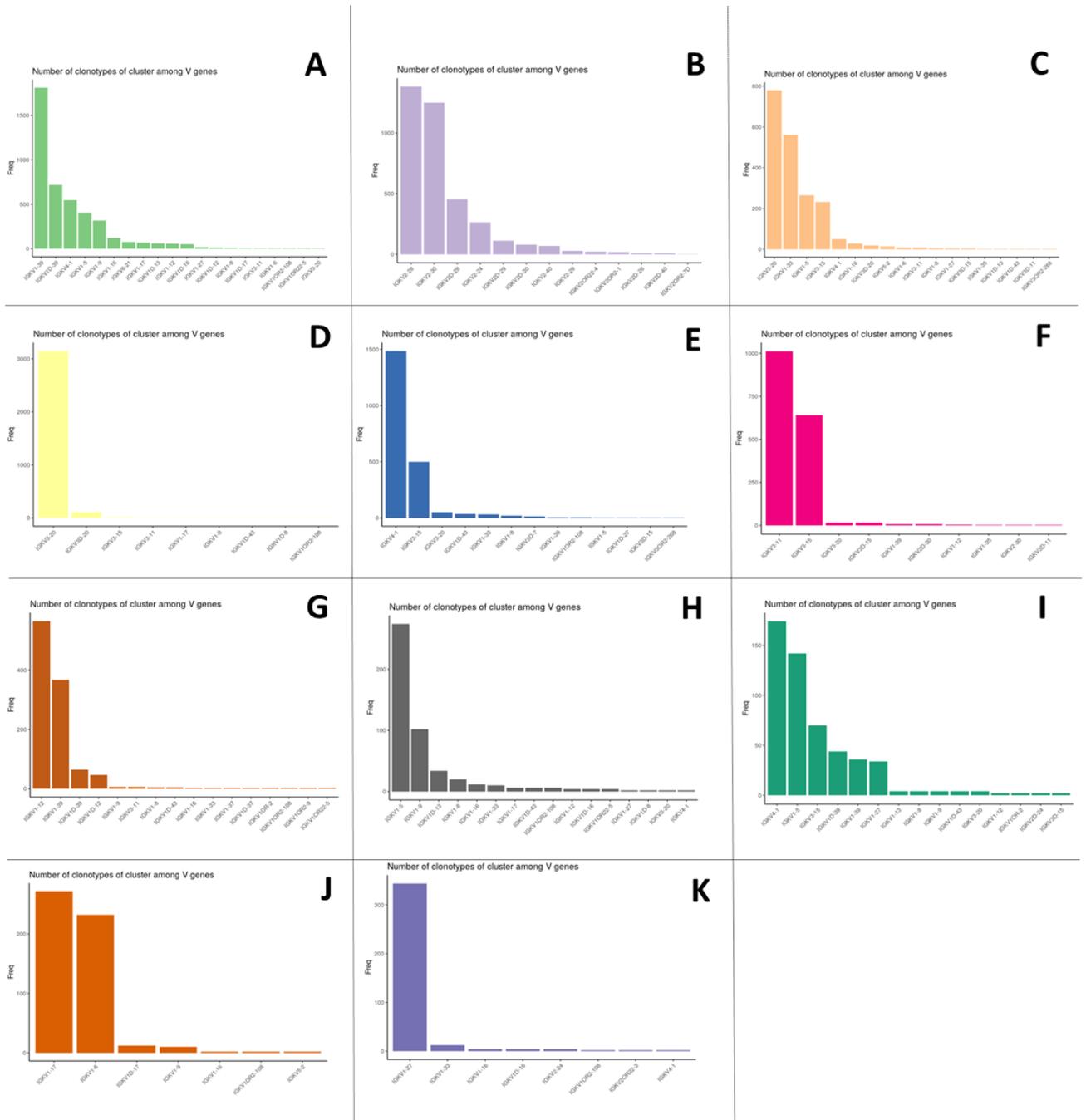
Supplementary Figure 6. Top enriched Reactome pathway in factor 1. Red : positive features; blue : negative features.



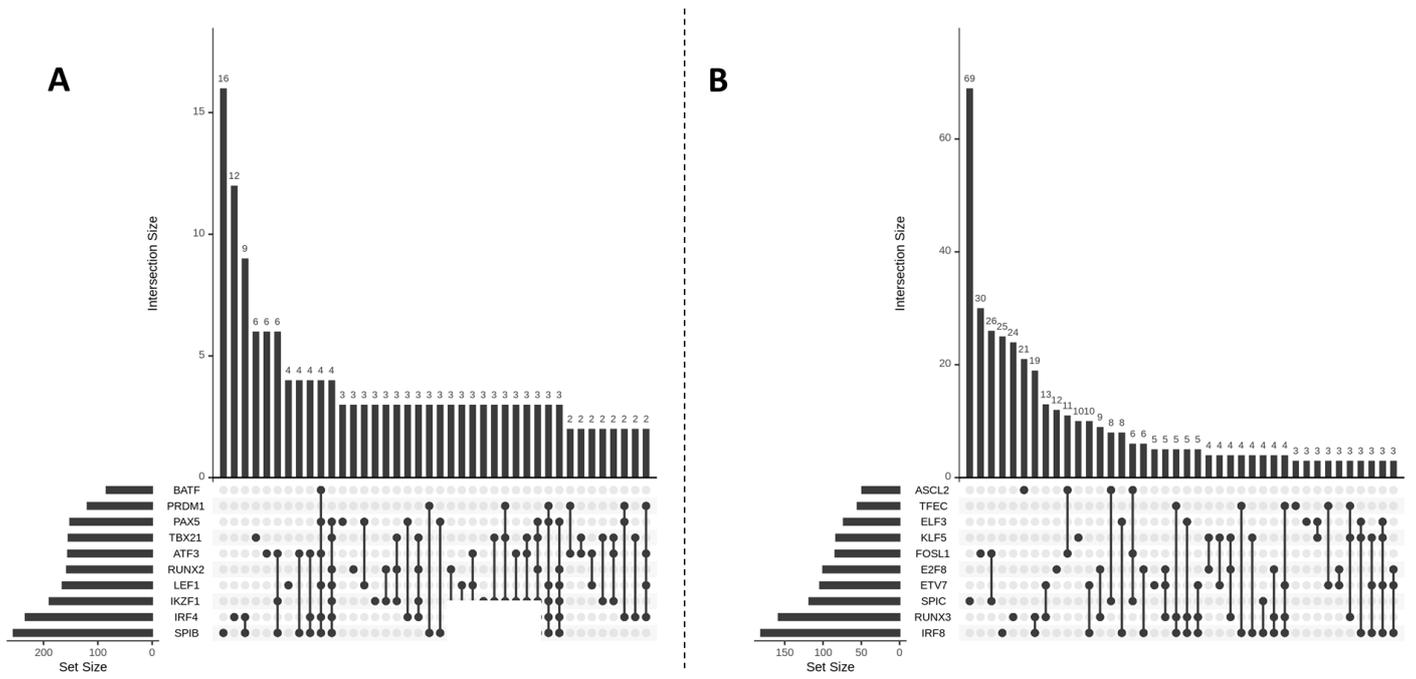
Supplementary Figure 7. Comparison between CTRL and CCC network. **(A)** Distribution of the correlation between paired features in one phenotype, when the correlation are high ($|r| \geq 0.8$) in the other phenotype. **(B)** Venn diagram of the conserved features in each network. **(C)** Venn diagram of the pairwise correlation between pairs of features in each network. Red : CCC; blue : CTRL.



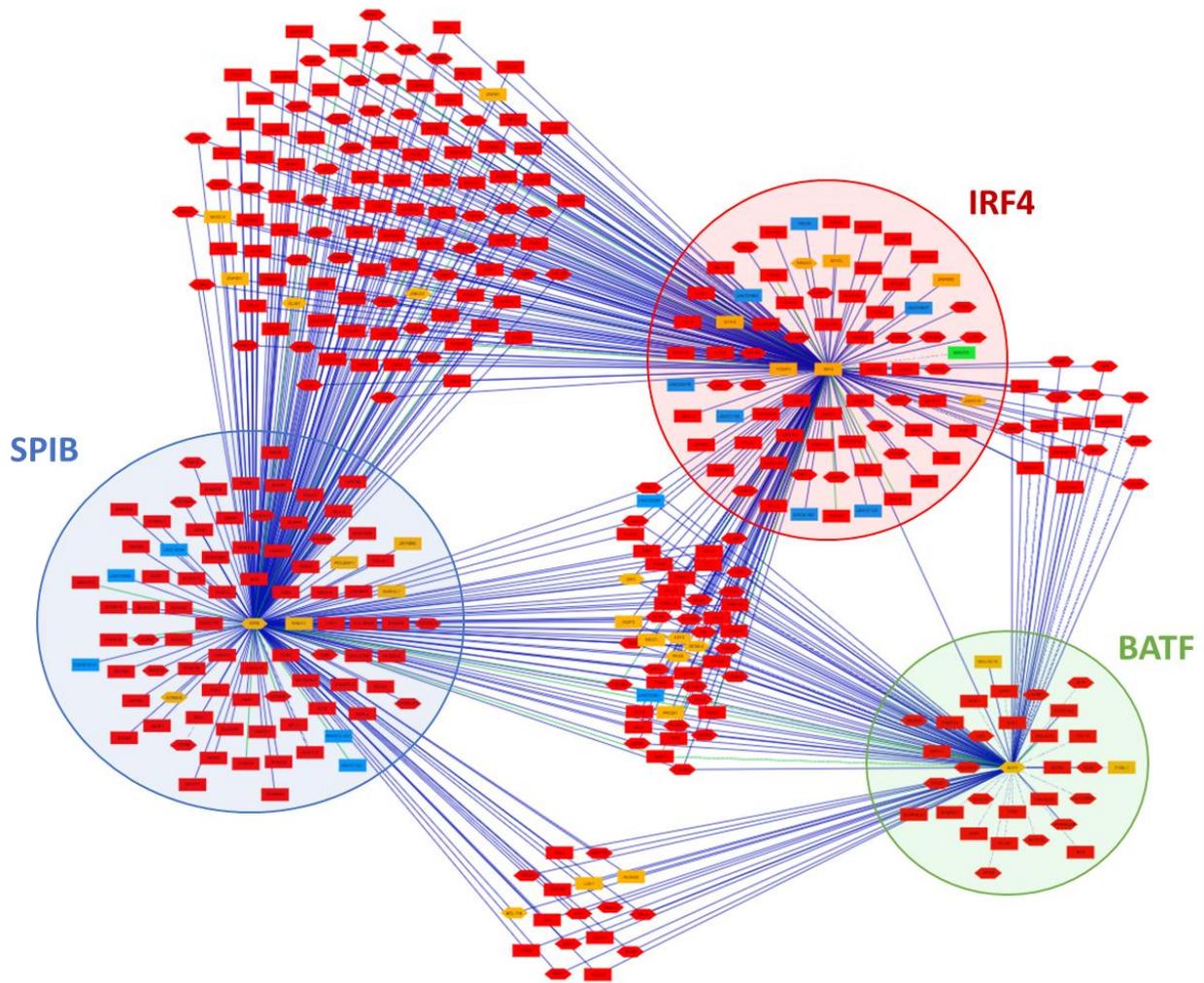
Supplementary Figure 8. Clonotypes analysis of T cell receptor. **(A)** Abundance of TCR clonotypes per phenotype. **(B)** Distribution of the number of samples sharing public clonotypes per phenotype : **1)** All samples, **2)** Control samples, **3)** CCC samples.



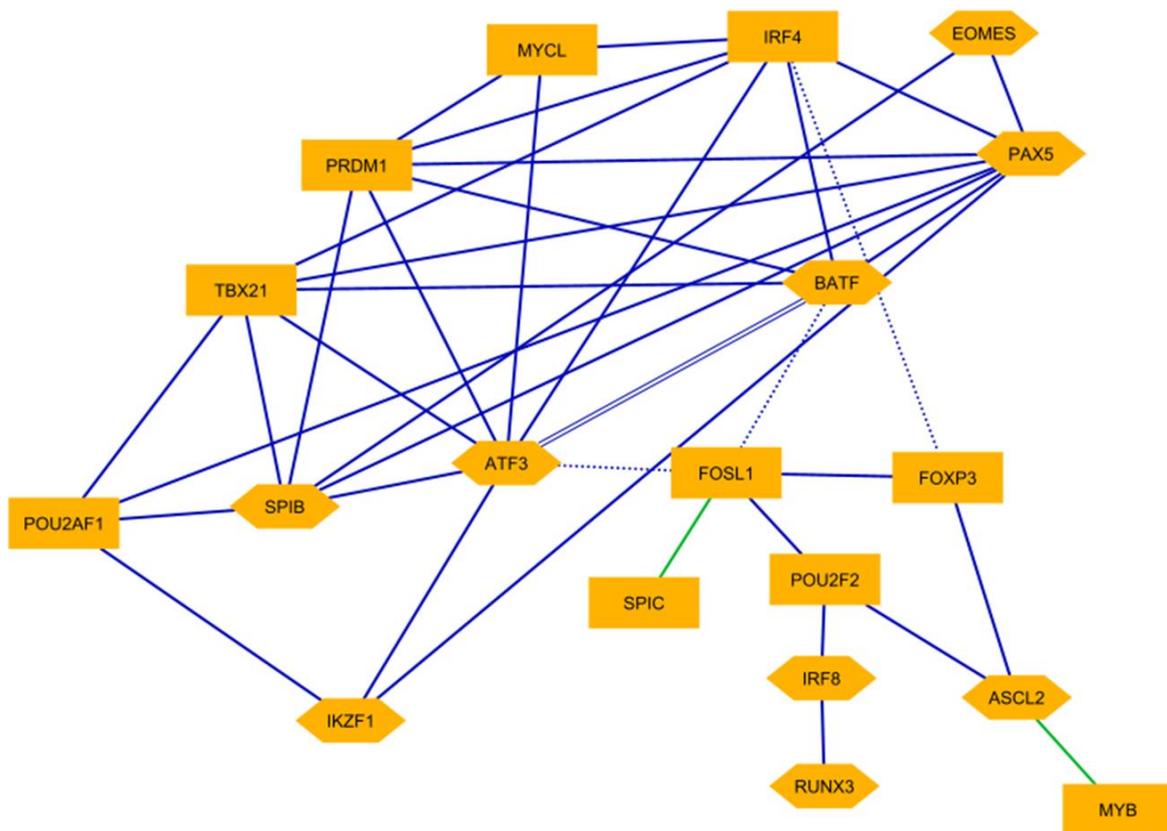
Supplementary Figure 9. Frequency distribution of the V genes associated with the BCR sequences. **(A)** Cluster 1, **(B)** Cluster 2, **(C)** Cluster 3, **(D)** Cluster 4, **(E)** Cluster 5, **(F)** Cluster 6, **(G)** Cluster 7, **(H)** Cluster 8, **(I)** Cluster 9, **(J)** Cluster 10, **(K)** Cluster 11.



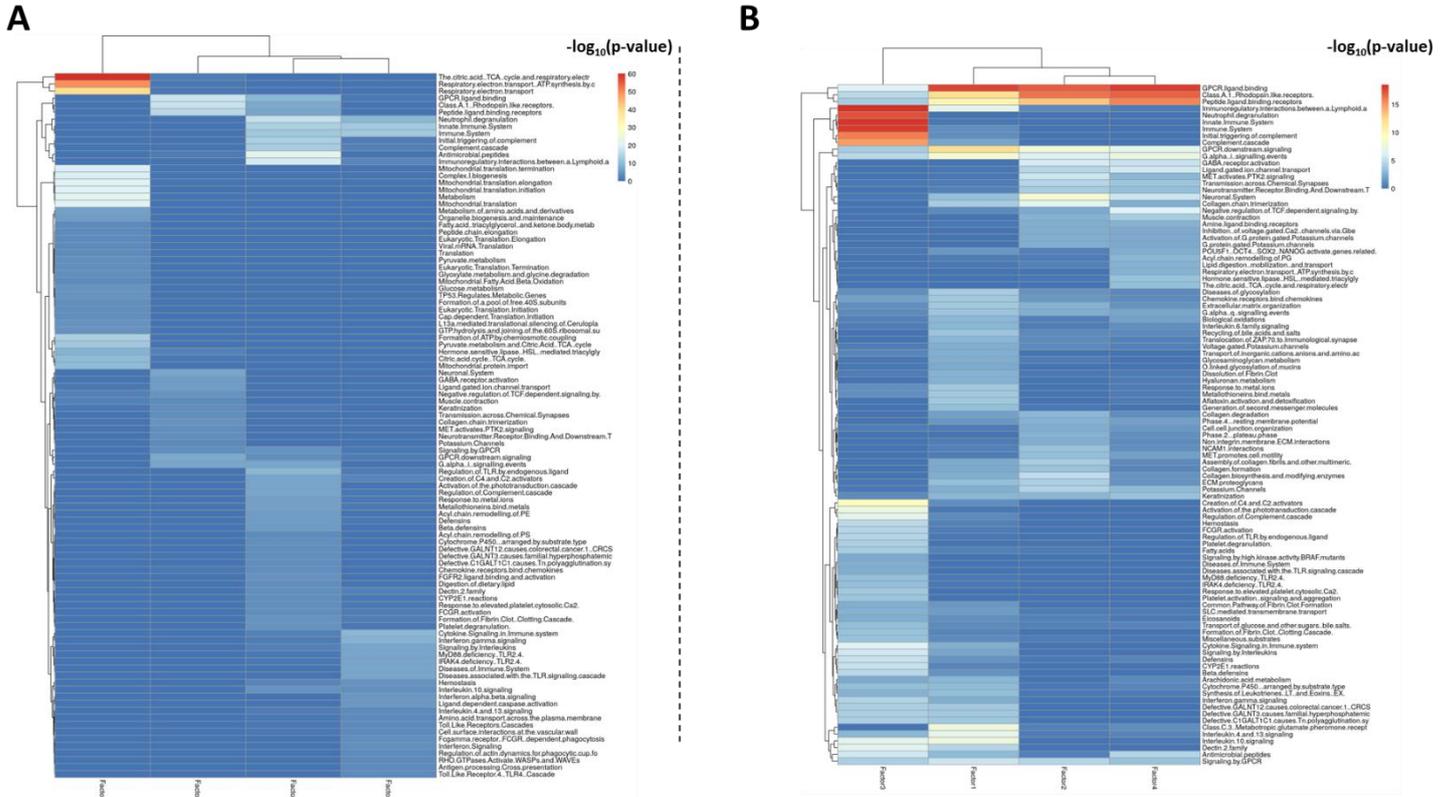
Supplementary Figure 10. Upset-plot of the top-10 connected TF in **(A)** cluster 2 and **(B)** clusters 1 and 3. The horizontal bars represent the total number correlated features to each TF, while the vertical bars represent the number correlated features unique to the different comparisons.



Supplementary Figure 11. Correlation network of IRF4 (red circle), SPIB (blue circle) and BATF (green circle) transcription factor targets genes. Nodes color: red: protein-coding genes, blue: lncRNA, green: miRNA, orange: transcription factor. Nodes shape: square: feature expression, hexagon : feature expression carrying at least one CpG selected with MOFA+ in its promoter. Edge line color: blue: positively correlated, green: negatively correlated. Edges line type: line: correlated in CCC network, dot: correlated in CTRL network, double line: correlated in both networks.



Supplementary Figure 12. Correlation network of the top-transcription factors identified in clusters 1, 2 and 3. Nodes color: orange: transcription factor. Nodes shape: square: feature expression, hexagon : feature expression carrying at least one CpG selected with MOFA+ in its promoter. Edge line color: blue: positively correlated, green: negatively correlated. Edges line type: line: correlated in CCC network, dot: correlated in CTRL network, double line: correlated in both networks.



Supplementary Figure 13. Reactome pathways significantly enriched in at least one MOFA+ CCC cluster.

Toutes les tables supplémentaires sont directement **accessibles en ligne** sur medRxiv.

DISCUSSION

La maladie de Chagas est une parasitose endémique d'Amérique Latine due au protozoaire *Trypanosoma cruzi*. L'évolution clinique de la maladie comprend une phase aiguë, généralement asymptomatique, et une phase chronique, la majorité des individus infectés demeurent asymptomatiques. Toutefois, des décennies après l'infection, des formes chroniques apparaissent chez 30% des individus, notamment cardiaques, appelées Cardiomyopathies Chagasiennes Chroniques (CCC). À ce stade, la parasitémie des patients atteints de CCC est très faible. De ce fait, bien que des traitements anti-parasitaires aient été développés, ils n'ont que peu d'efficacité contre cette pathologie, n'améliorant pas la fonction cardiaque. La CCC est l'une des cardiomyopathies les plus létales, et est la principale cause des décès dus à la maladie de Chagas, entraînant près de 10.000 morts par an.

Cette cardiomyopathie est caractérisée par une forte inflammation du myocarde avec un important infiltrat immunitaire, majoritairement composé de lymphocytes T, lymphocytes B et de macrophages. Comparés aux individus asymptomatiques (ASY), les patients atteints de CCC présentent un déséquilibre dans la balance des lymphocytes T helper (Th), avec une forte concentration de lymphocytes Th1 producteurs de cytokines pro-inflammatoires telles que l'IFN- γ et le TNF- α . À l'inverse, la production de cytokines anti-inflammatoires, et notamment d'IL-10, est plus faible chez les CCC. Les causes sous-jacentes de cette réaction inflammatoire exacerbée et non contrôlée demeurent méconnues. L'hypothèse d'une réaction auto-immune apparaissant chez les individus atteints de cette cardiomyopathie pourrait expliquer, tout du moins partiellement, l'activation continue de la réponse immunitaire.

Afin de comprendre le processus pathogénique associé aux CCC, des études omiques ont été menées ces dernières années. Bien que de nombreuses molécules provenant de différentes omiques aient été associées au risque de développer cette forme chronique, les interactions entre ces molécules ont été relativement peu explorées. C'est dans ce contexte que mon projet de thèse a été élaboré.

À partir de données génomiques, transcriptomiques et méthylomiques, nous avons mis en évidence la dérégulation de processus biologiques spécifiques de cette cardiomyopathie, ainsi que de nombreux facteurs génétiques et épigénétiques conduisant potentiellement aux altérations de l'expression génique observées. Majoritairement observées chez l'ensemble des patients atteints de CCC, nous avons également identifié des processus propres à des individus spécifiques, suggérant l'implication de processus pathogéniques différents conduisant tous à la destruction des tissus cardiaques.

1. Des différences génétiques et épigénétiques associées aux cardiomyopathies chagasiennes chroniques

1.1. Des gènes aux processus biologiques

1.1.1. Les gènes codants

Dans le cadre de mon projet, nous avons mené deux études de transcriptomique. La première a été réalisée sur des échantillons de tissu cardiaque de 10 patients atteints de CCC et 7 contrôles sains. L'expression des gènes a été quantifiée avec une puce à ADN, et nous a permis d'identifier 1.535 gènes différentiellement exprimés (DEG) entre ces deux phénotypes ([article 2](#)) (317). La deuxième étude transcriptomique a été réalisée en utilisant une partie de ces échantillons, incluant 6 contrôles sains et 8 CCC. La quantification de l'expression génique a été réalisée par séquençage de l'ARN (RNA sequencing), et a mis en évidence 1.407 DEG ([article 3](#)) (318). Ces analyses ont toutes deux identifié des dérégulations dans les processus immunitaires, majoritairement liés aux lymphocytes Th1 et à la réponse à l'IFN- γ . Dans une moindre échelle, les DEG identifiés sont également associés aux cellules dendritiques, cellules NK et lymphocytes B. La première étude a également relevé des altérations génétiques impliquées dans la contraction du cœur, l'hypertrophie, l'arythmie, le stress oxydatif et les processus mitochondriaux. Les différences fonctionnelles entre les deux études pourraient être liées aux techniques utilisées, ou aux bases de données servant à l'enrichissement fonctionnel (IPA Ingenuity dans le premier cas, Gene Ontology dans le second). Il est important de noter que des gènes associés à la réponse lymphocyte Th1 sont également retrouvés dans tous ces processus non immunitaires, appuyant leur rôle général dans la pathogénèse.

Le séquençage de l'ARN a également été mené sur 8 échantillons cardiaques de patients atteints de cardiomyopathies dilatées non inflammatoires (CMD). Seuls 9% des DEG identifiés entre les contrôles sains et les CMD sont en commun avec ceux associés aux CCC.

L'analyse des processus biologiques liés aux DEG montre que seuls les CCC présentent des perturbations dans les processus immunitaires, tandis que les CMD ne présentent pas ces altérations. Bien que ces deux types de cardiomyopathies partagent des perturbations liées aux muscles lisses et au transport des ions, le transport du calcium est principalement associé aux CCC. Une étude comparative des profils transcriptomiques des CCC et des CMD avait déjà été réalisée par notre équipe en 2005, et avait obtenu des résultats similaires, à l'exception du métabolisme du calcium, qui était alors majoritairement associé aux CMD et non aux CCC (240). En revanche, le métabolisme du calcium avait été majoritairement associé aux CMD, et non aux CCC. Le calcium joue un rôle essentiel lors de la reconnaissance du TCR par les antigènes, entraînant l'activation des gènes impliqués dans l'activation des lymphocytes T via la protéine NFAT (319). De plus, le calcium est également impliqué dans la contraction du muscle cardiaque, et une réduction de sa concentration a été associée à des cardiomyopathies dilatées (320). Par conséquent, bien que des processus liés au calcium aient été identifiés dans les deux pathologies, leur rôle dans la pathogenèse diffère. Le calcium est également impliqué dans la contraction du myocarde.

Tous ces résultats sont cohérents avec les connaissances actuelles sur les CCC et contribuent à une meilleure compréhension des processus sous-jacents au développement de cette maladie. Cependant, il est important de noter que l'étude transcriptomique seule ne permet pas de confirmer les altérations des processus identifiés. En effet, des modifications traductionnelles et post-traductionnelles peuvent avoir un impact significatif sur l'activité protéique, et donc sur les processus biologiques associés. Deux études protéomiques ont été menées sur des tissus cardiaques de patients atteints de CCC (259,321). Bien que le nombre de protéines étudiées soit drastiquement réduit comparé aux nombres de gènes obtenus par des études de transcriptomique, il est intéressant de noter que les altérations observées dans le fonctionnement mitochondrial (stress oxydatif, chaîne de transport d'électrons, production d'ATP) sont plus prédominantes en protéomique, suggérant des altérations protéiques qui ne pourraient pas être visibles au niveau transcriptomique.

1.1.2. Les ARN non codants

Les ARN non codants sont des ARN transcrits mais non traduits ayant un rôle dans la régulation de l'expression génique et protéique. L'implication de ces ARNs dans les CCC, et spécifiquement des micro ARN (miARN) et des ARN long non codants (ARNlnc), a déjà été démontrée dans la littérature (68,269).

Dans le cadre de mon projet, nous avons mené une analyse de l'expression des miARN à partir de tissu cardiaque de 4 contrôles sains et 8 CCC par RT-qPCR ([article 2](#)) (317). Cette étude a permis d'identifier 2 miARN significativement associés aux CCC (miR-146a et miR-155). MIR155HG, le gène codant pour miR-155, a également été identifié après séquençage de l'ARN ([article 3](#)) (318). Dans les myocardites virales, miR-155 est également surexprimé, et semble être impliqué dans la production de TNF- α , d'IFN- γ et d'IL-6, ainsi que dans l'infiltration des cellules immunitaires (322). L'absence de ce microARN semble diminuer la production d'IFN- γ et de TNF- α dans le tissu cardiaque de souris infectées par *T. cruzi* en phase aiguë (266).

De plus, 80 miARN tendent à être différentiellement exprimés entre les contrôles et les CCC, présentant tous une p-valeur significative avant correction multi-test. Parmi eux, miR-1, miR-133a et miR-133b ont déjà été associés aux CCC (68), suggérant un manque de puissance statistique dû au faible nombre d'échantillons. Ces 3 miARN sont à la fois sous-exprimés chez les CCC et les CDM comparés à des contrôles, mais également chez les CCC par rapport aux CMD. Connus pour leur implication dans des troubles cardiaques, il est possible qu'ils soient associés à la gravité des CCC par rapport à d'autres formes de cardiomyopathie.

Le séquençage de l'ARN, réalisé sur les tissus cardiaques de contrôles sains et de patients atteints de CCC ou de CDM, a fourni des informations sur les ARNlnc impliqués dans ces affections ([article 3](#)) (318). Nous avons identifié un total de 175 ARN non codants, parmi lesquels 145 étaient associés aux CCC. En outre, 143 de ces ARNlnc étaient spécifiques aux CCC et n'ont pas été détectés chez les patients atteints de CDM. Il est important de noter que les ARNlnc interagissent souvent avec les miARN en se liant préférentiellement à eux, ce qui les empêche de réguler leurs gènes cibles. Selon les bases de données publiques, MIAT, déjà connu pour son rôle dans les CCC, agit en ciblant le miARN miR-133a, que nous avons également identifié dans nos études (269). Il est important de noter que l'étude des ARN non codants seule ne fournit que peu d'informations sur les processus biologiques affectés, et nécessite des analyses approfondies pour comprendre leur impact sur la régulation de gènes cibles.

L'ensemble de ces études transcriptomiques a donc permis de révéler des processus biologiques d'intérêts associés aux CCC. De plus, en utilisant des méthodes de séquençage, fournissant des informations sur l'ensemble des transcrits du génome, ou des méthodes ciblées telle que la RT-PCR, elles ont permis une caractérisation approfondie des ARN non codants, pouvant avoir un impact sur la régulation génique. Cependant, les échantillons proviennent de patients atteints de CCC en stade terminal. De ce fait, les processus mécanistiques à l'origine des dérégulations géniques observées peuvent être difficiles à déceler, ou inopérants à ce stade de la maladie. L'accès au tissu cardiaque de patients ne peut se faire que dans un contexte de nécessité médicale, et il est donc impossible d'avoir accès à des échantillons asymptomatiques, ou atteints de CCC modérés. Pour pallier cette limite, des études ont été conduites sur le sang de patients asymptomatiques, atteints de CCC modérée ou atteints de CCC sévères (57). Bien que des processus spécifiquement cardiaques puissent ne pas être observables dans le sang, cette approche permet néanmoins d'identifier des marqueurs spécifiques des différents stades de la maladie.

1.2. L'implication de la méthylation dans le développement des cardiomyopathies chagasiennes chroniques

La méthylation de l'ADN est une modification épigénétique consistant à l'ajout d'un groupement méthyl sur une cytosine suivi d'une guanine, formant un îlot CpG. Ce changement structural peut impacter la fixation d'éléments régulateurs sur l'ADN, et donc indirectement réguler l'expression des gènes. Une étude menée par notre groupe avait permis d'associer des différences de méthylation aux CCC (260). Pour approfondir cette approche, nous avons conduit une analyse du méthylome en utilisant la méthode Illumina EPIC méthylation ([article 3](#)) (318). Cette technique repose sur l'utilisation de puces ciblant des sites spécifiques, ce qui nous a permis de quantifier les niveaux de méthylation à 850 000 positions, soit plus du double du nombre de sites que lors de notre première étude. De plus, chaque gène présent dans le génome est représenté au moins par une puce, offrant une vision globale du méthylome.

La comparaison du méthylome a été réalisée entre le tissu cardiaque de 6 contrôles sains et 8 patients atteints de CCC. Au total, 16.883 sites de méthylation ont été associés aux CCC, correspondant à 5.814 gènes.

Il est intéressant de noter que ces gènes sont impliqués dans des fonctions connues des CCC, telles que l'activation des lymphocytes, la production de cytokines ou le transport des ions. Cette comparaison souffre néanmoins de la même limite que les études transcriptomiques, comparant des individus sains à des patients atteints de CCC en stade terminal. Pour surmonter cette contrainte, nous avons également étudié le méthylome de patients asymptomatiques ou atteints de CCC modérés à partir d'échantillons sanguins. L'ensemble des sites différentiellement méthylés entre ces trois conditions a permis de distinguer les échantillons en fonction de leur phénotype avec une analyse en composante principale, démontrant leur implication dans les différents stades de la maladie. Bien que seuls 45% des sites de méthylation identifiés dans ces échantillons sanguins soient retrouvés dans les échantillons cardiaques, les gènes associés sont impliqués dans les mêmes fonctions biologiques, confirmant l'intérêt des analyses basées sur des données sanguines. Par ailleurs, un nombre limité de ces sites de méthylation a été utilisé pour prédire le phénotype des individus à l'aide de techniques d'apprentissage automatique, ce qui confirme leur importance dans le développement de cette pathologie ([article 4](#)) (323).

Le niveau de méthylation de l'ADN par lui-même ne permet pas de tirer des conclusions sur l'expression des gènes ou la fonctionnalité des protéines impliquées dans les processus biologiques d'intérêt. Cependant, il est important de noter que la méthylation de l'ADN peut avoir un impact significatif sur la régulation des gènes (324). Nos analyses ont révélé des variations significatives de méthylation dans des gènes liés aux processus biologiques associés aux CCC, ce qui renforce l'importance de la méthylation dans cette pathologie. Cela souligne également la nécessité d'approfondir nos recherches pour mieux comprendre les liens entre la méthylation de l'ADN et l'expression de gènes d'intérêt.

1.3. La génétique mitochondriale dans le développement des cardiomyopathies chagasiques chroniques

Compte tenu des altérations des processus mitochondriaux observés dans la maladie de Chagas, nous avons mis en place une étude comparant les mutations présentes dans les mitochondries cardiaques de 6 patients sains et de 36 patients atteints de CCC ([article 1](#)). Principalement, cette analyse nous a permis d'identifier une association entre les haplogroupes (ascendance ethnique commune) africains et européens avec le développement des CCC.

Au vu de la faible taille de cette cohorte, nous avons effectué une approche similaire à partir des mitochondries présentes dans le sang de 38 patients asymptomatiques, 34 atteints de CCC modérée et 36 atteints de CCC sévère. De plus, nous avons inclus des données publiques de Whole Exome Sequencing provenant de 289 individus sains issus de São Paulo. Cette analyse a montré qu'un nombre significatif de patients d'ascendance africaine étaient atteints de la maladie de Chagas et que, à l'inverse, la majorité des individus sains était d'ascendance européenne. Plusieurs macro-haplogroupes africains ont présenté des tendances d'association avec différents stades de la maladie (L0 plus fréquent chez les ASY et L3 chez les patients atteints de CCC). Ces résultats pourraient être associés à la prédisposition des individus d'ascendance africaine à être davantage exposés au parasite en raison de leur statut socio-économique moins favorable. À l'inverse, le macro-haplogroupe européen H a été associé au risque de développer des CCC. Des études présentes ont mis en évidence l'association entre les patients ayant l'haplogroupe H avec le risque de développer des cardiomyopathies (325). De plus, cet haplogroupe est associé à une augmentation des fonctions spécifiquement induite par la mitochondrie (326).

L'analyse des mutations mitochondriales a permis d'associer 712 variants au risque de développer une des formes de la maladie, dont 84 associés à des changements de la séquence protéique et annotés comme pathogéniques dans différentes bases de données publiques. Cependant, aucune de ces mutations n'est spécifique à un des phénotypes étudiés. Il est possible que les patients atteints de CCC ne partagent pas nécessairement de mutations mitochondriales communes, mais que différentes mutations conduisent au même phénotype. L'analyse des haplogroupes appuie cette idée, avec des tendances d'association de macro-haplogroupes, supportés par différentes mutations, au risque de développer des CCC. Cette hypothèse est supportée par une étude récente menée par notre équipe, qui a identifié des variants mitochondriaux famille-spécifique associés au risque de développer des CCC (235). Au-delà des mutations mitochondriales, cette hypothèse pourrait expliquer le faible nombre de mutations identifiées à grande échelle lors de différentes analyses GWAS et méta-analyses (232–234). Pour confirmer cette hypothèse, des analyses menées sur des lignées cellulaires porteuses de différentes mutations d'intérêt seraient intéressantes. Nous avons déjà mis en évidence l'impact de la stimulation de cardiomyocytes (AC-16) par de l'IFN- γ et du TNF- α sur la transcription de gènes mitochondriaux ([annexe 1](#)) (327).

L'introduction de différentes mutations dans ces cellules permettrait d'étudier les dérégulations génétiques propres à chacune, et leurs potentielles contributions indépendantes aux CCC. Des analyses similaires sur des lignées de souris seraient envisageables. Enfin, comme initié par notre équipe, l'analyse du génome de familles présentant des cas de CCC pourrait révéler différentes mutations spécifiques ayant pourtant des conséquences similaires. Il serait pour cela nécessaire d'accroître drastiquement le nombre de familles, ce qui peut se révéler difficile.

Les mutations mitochondriales en elles-mêmes peuvent ne pas suffire pour fournir une explication, au moins partielle, à la manifestation des CCC. Ces mutations s'inscrivent au sein de réseaux d'interactions complexes entre l'ADN nucléaire et l'ADN mitochondrial, modulant l'impact des différentes mutations (328). Il est donc possible que certaines mutations soient bien associées au risque de développer des CCC, mais que tous les individus porteurs ne manifestent pas de symptômes après infection par *T. cruzi*. Ainsi, différents facteurs externes pourraient influencer cette potentielle pénétrance incomplète, comme observé dans d'autres pathologies (329). Il est intéressant de noter que nous avons déjà associé un variant nucléaire à des dysfonctions mitochondriales en réponse à la stimulation par de l'IFN- γ chez des patients atteints de formes digestives de la maladie de Chagas ([annexe 2](#)) (330). Ces hypothèses et résultats appuient davantage la nécessité d'intégrer différentes sources de données, pouvant mettre en avant des processus complexes liés aux CCC.

1.4. Une vue d'ensemble des analyses omiques dans la compréhension des cardiomyopathies chagasiennes chroniques

Les analyses omiques menées pendant mon projet de thèse ont permis d'associer des gènes codants, des ARN non codants, des sites de méthylation et des mutations au processus pathogénique apparaissant chez les patients atteints de CCC. Bien qu'informatives, ces omiques seules ne constituent qu'une partie du processus pathogénique se développant dans les CCC. Afin d'avoir une vue exhaustive des différents éléments associés à cette pathologie, nous avons conduit une curation manuelle de l'ensemble des publications traitant de la maladie de Chagas ([article 5](#)) (331). Nous avons ainsi sélectionné 193 articles, contenant plus de 60.000 éléments associés à cette maladie à travers 5 espèces. Ces informations ont été rendues disponibles à la communauté au travers d'une base de données, ChagasDB, accessible directement en ligne.

Au-delà des omiques étudiées dans ce projet, elle recense des protéines, des molécules chimiques, des électrolytes et des hormones. Ces informations accentuent la complexité du processus pathogénique de la maladie de Chagas, à de nombreux niveaux biologiques.

A ce jour, seuls les composants biologiques observés *in vivo* et associés significativement à la maladie de Chagas ont été conservés. Cependant, de nombreuses études, principalement sur la souris, ont mis en place des techniques de “knock-out”, permettant d’estimer l’importance d’un gène ou d’une protéine dans la maladie. De plus, de nombreuses stimulations cellulaires *in vivo* ou *in vitro* par des cytokines ou des peptides de *T. cruzi* ont été réalisées. Ces différentes approches permettent une compréhension approfondie des réactions moléculaires spécifiques à la maladie de Chagas, et devront être ajoutées à ChagasDB dans une prochaine version.

1.5. Limites et perspectives des analyses simple omiques

Les analyses omiques effectuées dans ce projet ont identifié des caractéristiques génétiques et épigénétiques associées aux CCC, et confirmé l’implication de différents processus biologiques. L’impact de ces caractéristiques dans la régulation génique reste cependant à déterminer, ce qui nous a conduits à mettre en place des analyses multi-omiques.

Toutefois, toutes les analyses conduites sur du tissu cardiaque présentent deux limites majeures. Comme évoqué dans les paragraphes 1.1 et 1.2, ces études sont limitées par les échantillons obtenus par des patients atteints de CCC en stade terminal ou de contrôles sains. Par conséquent, les différents stades d’évolution de la maladie ne sont pas observables dans ces données. Une approche transcriptomique, menée sur le sang de patients ASY ou atteints de CCC modérée ou sévère a déjà été menée (57). Dans la continuité de cette étude, nous avons comparé les niveaux de méthylation et les profils génétiques mitochondriaux de patients présentant ces différents stades de la maladie. Cependant, ces différentes analyses n’ont pas été menées sur les mêmes patients, limitant la mise en place d’approches intégratives combinant ces données. De futures investigations devront être poursuivies dans cette direction afin de caractériser l’apparition des CCC, ainsi que leur évolution.

La deuxième limite de ces comparaisons est intrinsèquement liée à la réponse inflammatoire exacerbée se produisant dans le tissu cardiaque des patients atteints de CCC. En effet, un fort infiltrat immunitaire se produit dans le cœur de ces individus, qui est bien observable à l'échelle transcriptomique, comme démontré par notre analyse de déconvolution ([article 3](#)) (318). Ainsi, il convient de noter que les tissus cardiaques comparés présentent une composition cellulaire différente. Par conséquent, il est délicat de distinguer l'importance des gènes immunitaires dans cette pathologie spécifiquement, ces derniers pouvant être naturellement produits par les cellules immunitaires présentes dans les tissus. Néanmoins, nos expériences *in vitro* menées sur des cardiomyocytes (AC-16) stimulés par de l'IFN- γ et/ou du TNF- α ont mis en évidence des différences d'expression géniques significatives, notamment liées aux processus immunitaires. Cependant, ces variations pré et post-stimulation ont également été observées au niveau de gènes impliqués dans la réponse immunitaire, incluant plusieurs interleukines, TLR et gènes HLA, ainsi que des facteurs de transcription tels que TBX21, IRF1 ou IRF9 ([annexe 1](#)) (327). De ce fait, l'association directe de nombreux gènes liés à l'immunité avec les CCC ne semble pas uniquement attribuable à la présence de cellules immunitaire dans les tissus cardiaques. Les différentes analyses omiques menées sur ces tissus demeurent donc pertinentes, bien que limitées par le bruit associé à l'infiltrat immunitaire.

Une des solutions envisageables pour discerner les signaux spécifiques de types cellulaires précis serait de mener des approches à l'échelle cellulaire (single-cell). Il est en effet possible de mener de nombreuses analyses omiques à l'échelle de cellules uniques, incluant transcriptomique, épigénomique (méthylation et ouverture de la chromatine) et protéomique (332). Par exemple, une étude menée sur des échantillons cardiaques de patients atteints de cardiomyopathie dilatée et de contrôles sains a révélé des profils d'expression associés à la pathologie et spécifiques de différents types cellulaires, tels que les cardiomyocytes, les fibroblastes et les cellules endothéliales. De plus, les auteurs ont identifié différentes populations de fibroblastes et cellules endothéliales à travers les échantillons, certaines associées au développement de la cardiomyopathie (333). Une autre analyse a mis en évidence des lignées cellulaires en se basant sur leur ADN mitochondrial. Certains groupes de cellules, caractérisés par des profils génétiques mitochondriaux spécifiques, ont montré des variations d'expression génique. De plus, diverses mutations mitochondriales ont été associées à des altérations de la structure de la chromatine, entraînant probablement des régulations géniques distinctes.

Chez des patients atteints de leucémie lymphoïde chronique, l'identification de profils génétiques mitochondriaux similaires dans un sous-groupe de cellules myéloïdes (notamment des lymphocytes T et cellules NK) a conduit les auteurs à suggérer l'existence d'une cellule lymphoïde progénitrice commune, potentiellement associée au développement de ce type de cancer (334).

Ces différentes approches illustrent l'intérêt des analyses au niveau de la cellule unique, et l'apport que ces techniques pourraient avoir dans le cadre de la maladie de Chagas. Elles permettraient de caractériser les différents types cellulaires présents dans les échantillons cardiaques, et de potentiellement identifier des sous-populations particulières, comme cela a été avancé pour des lymphocytes T (335) ou B (128). De plus, elles approfondiraient les analyses des caractéristiques, en comparant leur expression par type (ou sous-type) cellulaire. Compte tenu du grand nombre de données publiques disponibles, notamment regroupées dans la base de données The human cell atlas (336), nous pourrions comparer les profils d'expression de différents types cellulaires associés à la maladie avec un grand nombre de contrôles (individus sains ou atteints d'autres maladies). Cela nous permettrait de mettre en évidence des expressions géniques inhabituelles par type cellulaire, et donc probablement liée à la pathologie.

2. Comprendre les régulations génétiques grâce à l'intégration multi-omiques

Le multi-omiques est une méthode d'analyse qui combine plusieurs types d'omiques afin d'étudier des processus biologiques complexes, identifier des biomarqueurs ou différencier plusieurs sous-groupes d'échantillons. Dans le cadre de mon projet de thèse, nous avons combiné différentes données omiques pour comprendre les causes des perturbations génétiques constatées chez les individus souffrant de CCC, et ainsi mieux comprendre les mécanismes sous-jacents de la réponse inflammatoire excessive et incontrôlée qui se manifeste dans leur tissu cardiaque. De ce fait, nous nous sommes reposés sur des échantillons cardiaques provenant de contrôles sains et de patients atteints de CCC sévère. Nous avons mis en place deux types d'approches : l'intégration tardive et l'intégration précoce.

2.1. L'intégration tardive

Dans le but d'identifier des éléments régulant les gènes différentiellement exprimés chez les patients atteints de CCC, nous avons effectué deux études reposant sur l'intégration de données transcriptomiques et miRnomiques, et de données transcriptomiques et méthylomiques. Dans le premier cas, nous nous sommes servis de la base de données IPA Ingenuity pour identifier des DEG ciblés par des miARN différentiellement exprimés ([article 2](#)) (317). Au total, 571 associations gène-miARN ont été sélectionnées, incluant 396 gènes et 67 miARN. Tous sont connus pour être associés à des processus biologiques d'intérêt, incluant l'inflammation, les gènes stimulés par l'IFN- γ , la fibrose et l'hypertrophie. 5 miARN (miR-125b-5p, miR-15a-5p, miR-296-5p, miR-29c-3p et miR-103a-3p) ciblent chacun plus de 20 gènes différentiellement exprimés, suggérant leur importance dans le processus pathogénique des CCC. Par ailleurs, ces 5 miARN sont connus pour être impliqués dans des processus conduisant au développement d'affections cardiaques. Plus généralement, un réseau de régulation contenant les miARN et les gènes cibles a été établi, en lien avec le développement d'arythmie, de myocardites, de fibroses et d'hypertrophies. L'implication majoritaire des miARN dans la régulation de gènes associés à des atteintes cardiaques est cohérente avec les résultats obtenus précédemment. Chez l'homme, les miARN différentiellement exprimés entre les patients atteints de CCC et ceux atteints de CDM ou des contrôles sains ont été prédits pour cibler des gènes différentiellement exprimés impliqués dans les maladies cardiovasculaires. Ces miARN sont par ailleurs sous-exprimés chez les CCC comparés à des contrôles sains (68). Dans un modèle murin, l'expression de 6 miARN a été significativement corrélée avec l'évolution de l'intervalle QT (mesure du système électrique cardiaque observable sur l'électrocardiogramme) suite à l'infection par *T. cruzi*. Ces derniers sont prédits comme ciblant des gènes impliqués dans la régulation du rythme cardiaque (266). Cette analyse nous a permis de caractériser l'impact des miARN dans les cardiomyopathies chagasiques. Cependant, malgré l'avantage de pouvoir associer des miARN à des gènes cibles grâce à l'utilisation de données publiques, cette approche ne permet pas une analyse approfondie des processus spécifiques à nos données. Par exemple, nous n'avons pas pu établir de corrélation significative entre l'expression des gènes et des miARN. Bien que les connaissances disponibles dans la littérature puissent renseigner sur des cibles potentielles, elles ne permettent qu'une étude partielle des processus de régulation qui se produisent chez les patients atteints de CCC.

Notre deuxième étude avait pour objectif d'estimer l'impact de la méthylation sur l'expression génique ([article 3](#)) (318). Pour ce faire, nous avons d'abord identifié les régions promotrices (TSS-1500, 5'UTR et premier exon) présentant des variations significatives de méthylation entre les individus en bonne santé et les patients atteints de cardiomyopathies chagasiennes chroniques. Ensuite, à l'aide de la base de données ReMap et de l'outil Ologram, nous avons prédit les facteurs de transcription qui se lient préférentiellement à ces régions. Enfin, en utilisant la base de données JASPAR et l'outil FIMO, nous avons restreint notre liste de facteurs de transcription à ceux dont le motif de fixation était significativement similaire à la séquence d'ADN de ces régions. Nous avons ainsi identifié 30 facteurs de transcription dont la fixation pourrait être influencée par les différences de méthylation. Tous ces facteurs de transcription sont impliqués dans des fonctions biologiques liées aux CCC, notamment la réponse lymphocytaire Th1, le système musculaire cardiaque et le système nerveux. Parmi eux, TBX21, RUNX3 et EBF1 ciblent le plus grand nombre de gènes. RUNX3 présente lui-même une région différenciellement méthylée dans son promoteur, potentiellement ciblée par 7 facteurs de transcription, incluant RUNX3. L'impact de la méthylation du promoteur de RUNX3 sur son expression a par ailleurs déjà été confirmé expérimentalement, appuyant la fiabilité de notre analyse (260). De plus, la coopération de RUNX3 et TBX21 a déjà été démontrée, induisant la production d'IFN- γ et l'inhibition d'IL-4 dans des cellules Th1 (337). Cette approche, plus complexe que celle mise en place dans l'intégration gène/miARN, s'est donc basée sur des données publiques et sur nos propres données pour associer la méthylation de l'ADN à l'expression génique. Elle a mis en avant un faible nombre de facteurs de transcription fortement impliqués dans la pathologie, contribuant à la compréhension du processus pathogénique. Les résultats obtenus nécessitent des confirmations expérimentales, comme des analyses en gène-rapporteur, pour caractériser l'impact de la méthylation à la fois sur l'expression du gène associé, mais également sur la fixation du (ou des) facteur de transcription(s) d'intérêt.

Toutefois, ces deux approches présentent plusieurs limites. D'une part, l'utilisation de bases de données publiques a permis d'associer différents types d'éléments. Ce faisant, elles ont restreint les analyses à des interactions connues dans la littérature, ne permettant pas l'identification d'éléments régulateurs encore méconnus. De plus, dans le cadre de l'analyse transcriptomique/méthylomique, seules les régions différenciellement méthylées situées en amont d'un gène ont été considérées, excluant les régions régulatrices intergéniques. D'autre part, la sélection des éléments différenciellement exprimés ou méthylés s'est faite indépendamment.

Bien que cette approche permette de choisir les caractéristiques les plus perturbées entre les phénotypes, elle ne prend pas en compte les covariations entre les caractéristiques des différentes omiques. En effet, à l'échelle d'une seule omique, les variations de certaines caractéristiques peuvent sembler faibles, mais avoir un impact significatif dans une autre omique. Enfin, ces études se sont limitées à l'intégration de deux omiques, ce qui a simplifié l'interprétation des résultats, mais n'a pas pris en compte les mécanismes plus complexes qui impliquent plusieurs caractéristiques différentes. C'est pour toutes ces raisons que nous avons mis en place par la suite une analyse plus complexe, intégrant directement 4 omiques différentes sans sélection des caractéristiques en amont.

2.2. L'intégration tardive

2.2.1. La complexité du pré-traitement des données

- La gestion des échantillons manquants

L'intégration multi-omiques a inclus des données de transcriptomique (séquençage de l'ARN), miRnomique (RT-qPCR), méthylomique (Illumina EPIC méthylation) et génomique (séquençage direct de l'ADN mitochondrial). Ces données ont été obtenues à partir du tissu cardiaque de 7 patients atteints de cardiomyopathies chagasiennes chroniques, et de 7 contrôles sains. Cependant, 1 échantillon contrôle et 4 échantillons CCC étaient absents dans au moins une omique. Au vu de la faible taille de la cohorte, il n'était pas envisageable de ne conserver que les échantillons présents dans toutes les omiques. Par conséquent, il était nécessaire d'imputer les données manquantes. Dans un premier temps, l'outil MI-MFA (298), dédié à l'imputation de données manquantes dans un jeu de données multi-omiques, a été employé. Toutefois, les échantillons imputés étaient très proches des échantillons non imputés, réduisant drastiquement la variabilité des données. Par conséquent, l'outil KNN-imputer de la librairie python sklearn a été utilisé. Ce dernier se sert des valeurs des k-plus proches voisins afin d'estimer les valeurs manquantes dans l'échantillon en question. Afin d'adapter cet outil à un jeu de données multi-omiques, une matrice de distance a été générée, en considérant le nombre d'échantillons manquants et le nombre de caractéristiques présentes dans chaque omique. Ainsi, chaque omique contribue équitablement à l'imputation, et le choix des k-plus proches voisins n'est pas influencé par un jeu de données spécifique.

L'imputation des échantillons manquants a été réalisée par phénotype pour s'assurer que les k-plus proches voisins soient des échantillons présentant des variations similaires dans le cadre de la pathologie. La validité de cette approche a été confirmée en réalisant plusieurs analyses multi-omiques complètes, dans lesquelles un échantillon existant a été enlevé aléatoirement. La majorité des caractéristiques identifiées comme dérégulées à l'échelle multi-omiques était commune à travers les itérations, appuyant l'efficacité de notre approche d'imputation.

- Le manque de puissance statistique

Malgré l'imputation des échantillons manquants, la faible taille de notre cohorte d'étude semblait constituer une limitation pour une analyse aussi complexe, ce constat ayant été corroboré par notre estimation de la puissance statistique grâce à l'outil MultiPower (301). Le nombre d'échantillons réels n'a pas pu être augmenté en raison des défis liés à l'obtention d'échantillons cardiaques humains, en particulier ceux provenant de sujets sains. Par conséquent, la création d'échantillons synthétiques s'est avérée inévitable. Ce constat explique également la décision d'imputer les échantillons manquants avant l'intégration multi-omiques. En effet, certaines méthodes d'intégration multi-omiques, comme MOFA (279), sont en mesure de gérer un nombre limité de données manquantes. Toutefois, en plus des défis liés à la génération d'échantillons synthétiques avec des données manquantes, le nombre d'échantillons présentant des données manquantes aurait été excessivement élevé après augmentation. Par ailleurs, nous n'avons sélectionné dans notre analyse que 7 patients atteints de CCC, alors qu'un plus grand nombre a été utilisé dans les différentes omiques. Cette restriction a été réalisée au regard du sexe des patients. En effet, tous les échantillons contrôles sont issus d'individus masculins, alors que la cohorte CCC est composée de patients masculins et féminins. Bien qu'aucune différence liée au sexe n'ait été relevée dans les analyses simple-omique, l'intégration multi-omiques préliminaire réalisée avec cette cohorte totale a révélé des caractéristiques spécifiques au sexe, introduisant du bruit dans notre analyse. Il est possible que ces observations soient directement liées à l'augmentation de données, augmentant la puissance statistique et relevant ainsi des spécificités d'une partie de la cohorte CCC.

SMOTE (304) est l'un des outils les plus connus pour générer des données artificielles, et est fréquemment employé dans l'apprentissage automatique pour effectuer de la classification d'échantillons.

Initialement prévu pour équilibrer des deux de données présentant une grande différence dans le nombre d'échantillons entre deux classes, SMOTE se sert des k-plus proches voisins d'un échantillon pour créer des échantillons synthétiques combinant les valeurs associées à deux échantillons réels proches. Toutefois, dans des données à grande échelle, cette méthode tend à réduire la variabilité globale du jeu de données, et induit des corrélations entre les échantillons (338). De ce fait, de nombreuses variantes de SMOTE ont été développées, dont KNNOR (307), qui utilise plusieurs échantillons proches ($n > 2$) pour créer des échantillons spécifiques. Bien qu'efficace dans une approche de classification, la pertinence de ces méthodes dans une étude reposant sur les co-variations de différentes caractéristiques n'a été que peu décrite. De ce fait, en utilisant de grandes cohortes d'échantillons provenant de la base de données TCGA, nous avons évalué la concordance de la structure des données synthétiques générées par ces deux méthodes par rapport aux données réelles en utilisant une Analyse en Composantes Principales (ACP). Nous avons également généré des échantillons synthétiques avec la technique de bootstrapping, consistant à copier un échantillon réel en lui ajoutant un faible poids. Ainsi, les différentes méthodes comparées synthétisent des échantillons à partir de 1, 2 ou plusieurs échantillons réels. Cette comparaison a montré que les échantillons générés avec SMOTE étaient les plus proches des échantillons réels.

Toutefois, comme discuté précédemment, la sélection des k-plus proches voisins peut être biaisée dans un jeu de données multi-omiques. Les précédentes analyses omiques réalisées avec nos données ont montré une homogénéité dans chacun de nos phénotypes, et nous ont permis d'opter pour une autre approche. Nous avons créé des échantillons synthétiques à partir de plusieurs échantillons réels, en considérant qu'un échantillon synthétique se situe au barycentre de ces échantillons réels. Nous avons modulé ce calcul en ajoutant un poids supérieur à un des échantillons réels. L'objectif est de garantir que l'échantillon synthétique suive la même distribution que les données d'origine. L'utilisation de plusieurs échantillons a également permis de limiter la création de corrélation entre deux échantillons réels générées par SMOTE. Cette approche ayant fourni des résultats similaires à SMOTE à partir des données TCGA, nous avons décidé de l'intégrer à notre analyse.

2.2.2. Le choix de la méthode

Il existe différentes méthodes d'intégration multi-omiques, reposant sur des principes mathématiques différents. Compte tenu de la faible taille de notre cohorte, les approches basées sur des méthodes d'apprentissages automatiques ont été exclues. Des méthodes basées sur les réseaux permettent d'intégrer directement des données publiques, assurant ainsi des associations entre features connues pour interagir ensemble. Parmi elles, iOmicsPASS (284) donne la possibilité à l'utilisateur de fournir ses propres données d'interactions. Cependant, le nombre et le type d'omiques traitées par l'outil est limité, et ne convenait pas à notre étude. Les méthodes basées sur la réduction de dimension permettent d'identifier à la fois des ensembles de caractéristiques et des ensembles d'échantillons présentant des variations communes. Ce type d'approches correspondait davantage aux objectifs de notre étude, et a donc été sélectionné. Plusieurs outils permettent ce type d'intégrations, de manière supervisée ou non supervisée. Les précédentes analyses omiques, et notamment l'analyse du génome mitochondrial, ont suggéré des spécificités propres à un (ou plusieurs) échantillon(s), mais pas nécessairement partagés par tous. De ce fait, l'utilisation de méthodes supervisées, cherchant les caractéristiques les plus discriminatives entre deux phénotypes, ne nous semblait pas adapté. Parmi les méthodes non supervisées, MCIA (287) et JIVE (286) ont été estimées comme les plus adaptées à nos données avec l'outil momix (313). Toutefois, la facilité d'utilisation et d'accessibilité des résultats offerte par MOFA nous a conduits à préférer cet outil. La comparaison des résultats avec ces trois outils a montré une grande convergence dans la sélection des caractéristiques à l'échelle multi-omiques, et donc un impact faible dans le choix de l'outil en particulier.

2.2.3. L'interprétation des résultats

L'analyse des facteurs obtenus avec MOFA a révélé un seul facteur (sur 7 au total) associé spécifiquement aux CCC. De plus, les autres facteurs semblent identifier des variations entre échantillons contrôles, mais pas entre les patients atteints de CCC. Les gènes associés au facteur 1 sont majoritairement impliqués dans la réponse immunitaire, mais présentent une diversité de fonctions (complément, signaux liés aux récepteurs B et T) qui n'étaient pas observées dans l'analyse seule du séquençage de l'ARN. De plus, des gènes associés aux processus mitochondriaux et cardiaques ont également été sélectionnés.

La sélection des caractéristiques propres au facteur 1 a permis d'identifier 1.214 gènes, 2.066 sites de méthylation et 58 miARN associés à la pathologie, dont près de 50% inconnus dans le cadre des CCC (d'après ChagasDB). L'analyse des réseaux de corrélations propres à chaque phénotype a révélé des différences de comportements évidents. Dans le but de comprendre l'impact des éléments régulateurs (facteur de transcription, miARN et ARNlnc) sur l'expression des gènes, nous avons effectué un clustering hiérarchique basé sur les corrélations entre ces différents types de caractéristiques. Nous avons ainsi défini trois clusters. Deux d'entre eux sont relatifs aux TLR et à la réponse immunitaire induite par le BCR, tandis que le troisième est associé aux lymphocytes T et à la réponse médiée par le TCR.

Au-delà des processus pro-inflammatoires déjà caractérisés dans les CCC, nous avons noté la présence de différents processus anti-inflammatoires, incluant notamment les facteurs de transcription IRF4 et SPIC. IRF4 est un facteur de transcription principalement produit par les lymphocytes B après la stimulation du BCR par les antigènes et le signal de costimulation avec les cellules T (339), mais également par les lymphocytes T après stimulation du TCR. Il est connu pour être crucial pour la différenciation des cellules Th2 et Treg (340), et médier le processus anti-inflammatoire après stimulation par FOXP3 (341). Dans nos données, IRF4 et FOXP3 ne sont corrélés que dans le réseau contrôle. Par ailleurs, dans le réseau de contrôles, il existe une corrélation entre IRF4 et l'expression de miR-375, et selon les prédictions de la base de données GRNdb, IRF4 est susceptible de réguler son expression. miR-375 a été décrit comme ayant un rôle dans la promotion des macrophages pro-inflammatoires M1 (342). De plus, ce miARN inhibe la voie PD1/PD-L1, jouant un rôle régulateur dans les lymphocytes T (342). L'absence de corrélation entre l'expression de FOXP3 et d'IRF4 chez les CCC pourrait suggérer une altération dans la stimulation d'IRF4, inhibant son rôle régulateur de miR-375 et son action anti-inflammatoire. Par ailleurs, l'action de IRF4 sur ses gènes cibles dépend du type cellulaire. En effet, dans les lymphocytes B, il interagit préférentiellement avec SPIB, et dans les lymphocytes T, avec BATF. Dans nos données, IRF4 est corrélé, dans le réseau CCC, avec CTLA4, un marqueur des lymphocytes Treg. De même, il est corrélé avec SOCS3, connu pour induire une réponse lymphocyte Th2 (343). Dans cette même idée, SPIC est corrélé avec les TLR6, 7 et 8, ainsi qu'avec IL-4, une cytokine promouvant la différenciation des lymphocytes CD4⁺ en Th2. SPIC est également connu pour mener une réponse anti-inflammatoire suite à son activation par les TLR (344).

Une seconde intégration multi-omiques a été menée sur avec les échantillons CCC spécifiquement. Elle a permis de mettre en évidence 4 facteurs discriminant 5 des 7 patients. Les gènes impliqués dans ces facteurs sont impliqués dans des processus biologiques associés à la maladie de Chagas, confirmant que les caractéristiques discriminant les patients sont dues à la pathologie, et non à des variables externes. La sélection des caractéristiques propres à chaque patient a mis en évidence des processus biologiques divergents, montrant que ces patients présentent des réponses moléculaires propres à la maladie. En revanche, aucun de ces processus n'est associé aux lymphocytes B. L'analyse détaillée des répertoires des lymphocytes B et T a montré une grande convergence dans les séquences des chaînes légères du BCR, démontrant une pression de sélection continue et commune à tous les patients atteints de CCC. L'ensemble de ces observations tendent à prouver que la réponse des lymphocytes B est commune à tous les patients, alors que des processus associés aux mitochondries, aux lymphocytes T ou aux TLR seraient variables selon les individus. Compte tenu de la faible parasitémie présente chez les patients atteints de CCC, l'hypothèse d'une réaction auto-immune, fortement appuyée dans la littérature (147), semble très probable. Des outils permettant d'estimer la probabilité de liaison épitopes/paratopes d'intérêt ont été développés, comme par exemple AbAdpt (345), mais ne sont pas applicables à nos données. En effet, il n'est pas possible d'associer les clonotypes des chaînes légères à ceux des chaînes lourdes, et donc de reconstituer *in silico* les différents BCR séquencés. Des analyses complémentaires seront nécessaires, comme des séquençages des BCR et TCR à l'échelle d'une cellule unique (346).

À la différence de l'intégration multi-omiques combinant les échantillons contrôles et les échantillons CCC, l'intégration CCC-spécifique a permis de sélectionner des variants mitochondriaux. Le patient PMG présente de fortes dérégulations dans les gènes associés aux processus mitochondriaux, et notamment au complexe I. Une des mutations associées à ce patient se situe dans le gène MT-ND1, situé dans le complexe I. Cette mutation est partagée par deux autres patients. Cependant, ces derniers ne présentent pas les altérations géniques associées à la mitochondrie observées chez PMG. Cette mutation pourrait être liée à une surexpression de deux miARN, miR-369-5p et miR-591. miR-369 a notamment été décrit comme impliqué dans la réduction de l'inflammation chez le rat (347), et à une réduction de la production de NO (348). Compte tenu des perturbations mitochondriales observées chez CCC6, l'expression de ce miARN pourrait être induite pour tenter de réduire les dommages mitochondriaux et la génération de stress oxydatif.

Le patients CCC3, quant à lui, présente des différences d'expression de caractéristiques précédemment associées à la réponse inflammatoire, et notamment du facteur de transcription IRF4. Ces résultats suggèrent des processus biologiques spécifiques à chaque patient, chacun résultant au développement de CCC. Ils renforcent l'hypothèse précédemment avancée, selon laquelle les patients présentent des profils génomiques uniques qui modifient de manière différentielle les processus biologiques liés à la pathogenèse de la maladie de Chagas, le tout contribuant au développement des CCC.

L'ensemble de ces analyses démontrent l'intérêt de l'intégration multi-omiques pour la compréhension des processus associés aux CCC, offrant ainsi des résultats prometteurs. Les réseaux que nous avons construits ont servi de cadre biologique et ont permis d'identifier des éléments de régulation intéressants. Cependant, il est important de noter que ces réseaux sont principalement basés sur des corrélations filtrées à partir d'interactions potentielles répertoriées dans diverses bases de données publiques, dont la plupart sont des prédictions. Cette méthode de construction conduit à des réseaux probables, bien que fortement interconnectés et parfois ambigus. Il est envisageable que certaines caractéristiques d'intérêt aient de nombreuses interactions dans le réseau, mais qu'en réalité, leur nombre soit limité, certaines de ces interactions étant corrélées mais n'ayant qu'un impact indirect. L'analyse de ces réseaux potentiels est une tâche complexe qui demande un investissement en temps considérable pour l'interprétation des résultats, et des analyses supplémentaires sont actuellement en cours.

2.2.4. Limites et perspectives

Malgré l'intérêt des résultats obtenus dans cette analyse, elle présente plusieurs limites. Comme détaillé dans le paragraphe 1.5, l'infiltrat immunitaire présent dans le tissu cardiaque des patients atteints de CCC induit une différence de profils cellulaires entre les contrôles et ces patients. Cette différence est d'autant plus marquée que nous avons identifié des facteurs de transcriptions d'intérêt, impliqués dans la réponse anti-inflammatoire, mais dont les cibles sont associées à la fois à l'inflammation et à sa répression. Leur impact sur l'expression génique peut être type cellulaire-dépendant, et il est indispensable d'effectuer des analyses à l'échelle de cellules uniques pour confirmer, ou infirmer, leur intérêt dans la pathologie. Des méthodes d'intégrations multi-omiques à l'échelle de cellules uniques ont vu le jour. Parmi elles, la deuxième version de MOFA, MOFA+, offre la possibilité d'intégrer des données de cellules uniques.

Ces approches permettraient de caractériser les régulations spécifiques à différents types cellulaires et, dans le cadre des CCC, de potentiellement identifier des caractéristiques impliquées dans la promotion des lymphocytes T en Th1, ou à l'inverse dans l'absence de production de lymphocytes Treg. L'application de ces méthodes sur des échantillons sanguins de patients asymptomatiques, atteints de CCC modérée ou de CCC sévère conduirait à un suivi du développement des CCC au niveau moléculaire et cellulaire, et à une compréhension approfondie du processus pathogénique.

Cependant, les approches à l'échelle de la cellule unique apportent un niveau de complexité supérieur que celles menées sur l'ensemble des cellules d'un tissu, et cette complexité est d'autant plus importante au niveau multi-omiques. Bien que ces approches soient indispensables pour approfondir nos analyses, nous avons dans un premier temps pallié au problème d'hétérogénéité cellulaire en menant une analyse spécifiquement sur les patients atteints de CCC. Elle a permis de différencier 5 des 7 patients, et d'associer des caractéristiques spécifiques à chacun d'entre eux. Cependant, seuls 3 d'entre eux manifestent des fonctions biologiques spécifiques. Concernant les 4 autres, les profils transcriptomiques de 3 ont été imputés, car ils étaient absents des données de séquençage de l'ARN. L'imputation ayant été réalisée à partir des k-plus proches voisins, et le nombre de patients atteints de CCC étant de 7, il n'est pas surprenant de ne pas détecter de variations génétiques spécifiques dans ces échantillons. Les méthodes de pré-traitement des données employées, incluant l'imputation et la génération d'échantillons synthétiques, ont permis de mener cette analyse et d'obtenir des résultats cohérents et pertinents au vu des connaissances de la pathologie. Toutefois, ces méthodes ont également réduit la variabilité des données, et présentent des limites dans l'analyse détaillée de chaque échantillon. L'obtention d'une plus grande cohorte avec un nombre limité de données manquantes permettrait d'améliorer et d'approfondir ces analyses.

Une contrainte supplémentaire à noter dans cette étude est le nombre restreint d'omiques disponibles. En effet, de nombreuses molécules, incluant des protéines, mais également des hormones ou des composés chimiques, ont été associées aux CCC ([article 5](#)) (331). Par conséquent, les résultats obtenus ne fournissent une vue que partielle des dérégulations induites dans le cadre de la maladie. L'intégration de nouvelles omiques conduirait à une étude plus précise et plus approfondie des interactions entre les différentes omiques, et à la mise en place du processus pathogénique conduisant aux cardiomyopathies chagasiques. Enfin, il est important de noter que tous les résultats ont été obtenus *in silico*.

Bien qu'intéressants, ils demeurent des prédictions. L'identification d'interactions prometteuses entre différentes caractéristiques contribuant potentiellement au développement des CCC devra être validée expérimentalement, par le biais de techniques telles que l'utilisation de gènes rapporteurs.

3. Contribution de cette thèse dans la recherche pour la maladie de Chagas

Au cours de ce projet de thèse, nous avons conduit plusieurs études au niveau simple, bi- et multi-omiques. Bien qu'elles comportent diverses limites et exigent des investigations supplémentaires, elles ont tout de même apporté des connaissances non négligeables sur la maladie de Chagas, avec de potentiels impacts à court et long termes.

3.1. Des progressions dans la compréhension du processus pathogénique associé aux cardiomyopathies chagasiques chroniques

Les premières études menées lors de ce projet ont été conduites sur des données de transcriptomiques. Elles ont permis d'identifier des processus biologiques affectés dans les cardiomyopathies chagasiques chroniques, en comparant les profils d'expression de patients atteints de CCC avec des patients atteints de CMD ou des contrôles sains. Cependant, plusieurs études précédentes ayant mené des travaux similaires, les processus biologiques impliqués dans les CCC étaient déjà bien connus. De même, l'identification d'haplogroupes mitochondriaux susceptibles de favoriser le développement des CCC est intéressante et dans la continuité des analyses menées sur la génétique mitochondriale, mais n'apporte que peu de nouveauté dans la compréhension du processus pathogénique.

L'analyse de caractéristiques épigénomiques, telles que les sites de méthylation ou les ARN non codants, a fourni un apport plus important dans les connaissances relatives à cette maladie. Concernant les miARN, des études antérieures avaient déjà été effectuées sur ce type de données, et peu de nouveaux éléments ont été identifiés. En revanche, nous avons mené pour la première fois une caractérisation avancée des ARNinc dérégulés dans les CCC, au regard de contrôles sains mais également de CDM.

L'impact de la méthylation de l'ADN dans les CCC avait déjà été rapporté précédemment, mais nous avons conduit une étude à plus grande échelle, et identifié de nombreux gènes potentiellement affectés par des différences de méthylation. De plus, nous avons pour la première fois conduit une étude approfondie de la méthylation sur le sang de patients à différents stades de la maladie. L'implication des gènes différentiellement méthylés dans des processus associés aux CCC tend à confirmer, d'une part, l'impact de la méthylation dans le développement de cette pathologie et, d'autre part, la possibilité de mener des études omiques pertinentes sur le sang de patients à différents stades de la maladie.

Au-delà de nos analyses personnelles, la construction de la base de données ChagasDB pourrait avoir un impact important dans la recherche sur la maladie de Chagas. Avec le développement des techniques d'analyses à grande échelle, de nombreuses études omiques ont vu le jour, fournissant des ensembles de caractéristiques associées à la pathologie. Pour autant, l'intégration de ces différents travaux peut parfois être complexe au regard du nombre important de caractéristiques identifiées. Cette base de données pourrait devenir une référence pour de nombreuses études, permettant d'associer de futurs résultats à différentes formes de la maladie ou à différentes souches du parasite, ainsi que d'évaluer leur nouveauté.

Le plus grand apport de l'ensemble de ce projet est évidemment l'intégration de plusieurs omiques. Qu'elle ait été faite de manière tardive ou précoce, la combinaison de différentes caractéristiques a permis de mettre en avant des interactions potentielles intéressantes, caractérisant partiellement le processus pathogénique mis en place dans les CCC. Alors que les analyses d'intégration tardives se sont concentrées sur un nombre restreint de possibles régulations, l'intégration précoce a fourni une quantité considérable de potentiels mécanismes régulateurs directement observables au travers de réseaux de corrélations. Bien que différentes limites techniques restreignent leur interprétation, ils pourront être exploités lors de futures analyses précises, fournissant plusieurs pistes expliquant partiellement le développement des CCC.

L'ensemble de ces études a donc conduit à des avancées directes dans la compréhension de la pathologie, à travers l'identification de cibles particulières, mais fournit également de nombreux supports pour des analyses futures, notamment avec le développement de ChagasDB, et la mise à disposition de réseaux d'interactions potentielles à l'échelle multi-omiques.

3.2. Au-delà de la maladie de Chagas : des approches bio-informatiques avancées

Ce projet de thèse reposait sur la compréhension du processus pathogénique associé au développement des CCC. Pour ce faire, différentes approches d'intégration ont été développées. En ce qui concerne l'intégration tardive, le protocole technique employé est relativement commun. Il est intéressant de noter que l'utilisation de diverses données publiques, telles que les données de Chip-Seq provenant de la base de données ReMap ou les motifs de fixation de facteurs de transcription de la base de données Jaspar, a renforcé et consolidé les résultats obtenus.

En revanche, le protocole d'intégration précoce élaboré dans notre approche multi-omique présente différents aspects techniques non négligeables, en tenant compte des études antérieures menées dans la littérature. Nous étions confrontés à un jeu de données multi-omiques avec une très faible cohorte d'échantillons, dont certains étaient manquants dans une ou plusieurs omiques. Des méthodes complexes, reposant sur des méthodes d'apprentissage automatique, ont déjà été développées pour pallier à ces limites, avec des résultats très satisfaisants. Toutefois, elles nécessitent un jeu de données d'entrée relativement grand, et ne semblaient pas adaptées à notre cohorte d'étude. Nous avons exploré différentes méthodes d'imputation et d'augmentation de données, mais qui n'ont pas été convaincantes. L'approche que nous avons choisi de mener nous a semblé, après avoir été analysée et qualifiée de plusieurs manières, être de qualité satisfaisante et nous a fourni des résultats cohérents. Le développement de ce protocole a été réalisé pour le traitement de nos données, et non pas pour devenir une référence dans le traitement de données multi-omiques. Néanmoins, de nombreuses études se retrouvent confrontées à des problèmes techniques similaires à ceux que nous avons rencontrés, et notre approche pourrait servir de base à l'intégration de ces données dans des situations similaires aux nôtres.

Au-delà du pré-traitement des données, nous avons également exploité les résultats d'intégration multi-omiques fournis par MOFA+ afin d'obtenir des réseaux de corrélation cohérents avec les connaissances publiques. À l'heure actuelle, peu d'études ont mené une caractérisation approfondie des données multi-omiques dans la compréhension d'un phénomène d'intérêt. La quantité de résultats obtenus pourrait en être une des causes.

En créant différents réseaux de corrélation par condition, en les filtrant avec de nombreuses bases de données et en les comparant, nous avons mené une étude approfondie des co-variations entre caractéristiques observées au niveau multi-omiques. Malgré la complexité d'interprétation de ces réseaux, cette approche pourrait fournir des pistes d'analyses de données multi-omiques pour d'autres études ayant des objectifs similaires aux nôtres.

3.3. De l'analyse *in silico* à une application médicale

L'ensemble des observations réalisées au niveau simple ou multi-omiques au cours de mon projet de thèse pourraient conduire à l'identification de cibles thérapeutiques, régulant la réponse inflammatoire et, de ce fait, limitant les dommages occasionnés au tissu cardiaque. Pour ce faire, une détection précoce des formes cardiaques est primordiale. Notre analyse menée sur des sites de méthylation obtenus dans le sang de patients asymptomatiques ou atteints de CCC modérée ou sévère ont permis de prédire à 95% l'apparition des formes cardiaques, et à 96% la sévérité des CCC ([article 4](#)) (323). Ces résultats doivent être confirmés sur une plus grande cohorte, mais ils offrent une piste intéressante dans la détection des différentes formes de la maladie. Bien que la mesure de la méthylation de l'ADN soit encore coûteuse à ce jour, l'implication des gènes portant ces sites de méthylation dans des processus connus de la pathologie tend à confirmer leur importance. Des mesures de leur expression dans le sang de patients atteints de différentes formes de la maladie de Chagas pourrait être conduite afin d'étudier leur potentiel intérêt en tant que biomarqueurs.

De plus, l'identification de processus biologiques spécifiques à différents patients atteints de CCC tendrait à démontrer des causes divergentes dans le développement de cette pathologie. Cette analyse a notamment été possible grâce à l'augmentation de données, qui a amplifié les signaux patients-spécifiques. Ces résultats doivent néanmoins être confirmés à plus grande échelle. Dans le cadre du cancer, de nombreuses recherches tendent à prouver l'intérêt de la médecine de précision, considérant les différences de réactions aux traitements des patients (349). La caractérisation de différents processus propres à des patients atteints de CCC pourrait conduire à des traitements spécifiques, agissant à différents niveaux mais conduisant tous à une réduction de la réponse inflammatoire exacerbée dans les tissus cardiaques.

REFERENCES

1. Lidani KCF, Andrade FA, Bavia L, Damasceno FS, Beltrame MH, Messias-Reason IJ, et al. Chagas Disease: From Discovery to a Worldwide Health Problem. *Front Public Health*. 2 juill 2019;7:166.
2. Chao C, Leone JL, Vigliano CA. Chagas disease: Historic perspective. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*. 1 mai 2020;1866(5):165689.
3. Lewinsohn R. Carlos Chagas and the discovery of Chagas' disease (American trypanosomiasis). *Journal of the Royal Society of Medicine*. juin 1981;74(6):451.
4. Bestetti RB, Restini CBA, Couto LB. Carlos Chagas Discoveries as a Drop Back to Scientific Construction of Chronic Chagas Heart Disease. *Arq Bras Cardiol*. juill 2016;107(1):63-70.
5. Chagas C, Villela E. Cardiac Form of American Trypanosomiasis. *Memorias do Instituto Oswaldo Cruz* [Internet]. 1922 [cité 4 oct 2023];14(1). Disponible sur: <https://www.cabdirect.org/cabdirect/abstract/19232900676>
6. Bestetti RB, Martins CA, Cardinalli-Neto A. Justice where justice is due: A posthumous Nobel Prize to Carlos Chagas (1879–1934), the discoverer of American Trypanosomiasis (Chagas' disease). *International Journal of Cardiology*. 1 mai 2009;134(1):9-16.
7. Andrews KT, Fisher G, Skinner-Adams TS. Drug repurposing and human parasitic protozoan diseases. *Int J Parasitol Drugs Drug Resist*. 24 mars 2014;4(2):95-111.
8. Cavallo MJ, Amelotti I, Abraham L, Cueto G, Gorla DE. Rural houses infestation by *Triatoma infestans* in northwestern Argentina: Vector control in a high spatial heterogeneous infestation area. *PLoS One*. 2 août 2018;13(8):e0201391.
9. Jansen AM, Xavier SC das C, Roque ALR. *Trypanosoma cruzi* transmission in the wild and its most important reservoir hosts in Brazil. *Parasit Vectors*. 6 sept 2018;11:502.
10. Fontes A, Santos BS, Chaves CR, Figueiredo RCBQ, Fontes A, Santos BS, et al. II-VI Quantum Dots as Fluorescent Probes for Studying Trypanosomatides. In: *Quantum Dots - A Variety of New Applications* [Internet]. IntechOpen; 2012 [cité 4 oct 2023]. Disponible sur: <https://www.intechopen.com/chapters/34822>
11. Meiser CK, Schaub GA. Xenodiagnosis. In: Mehlhorn H, éditeur. *Nature Helps.: How Plants and Other Organisms Contribute to Solve Health Problems* [Internet]. Berlin, Heidelberg: Springer; 2011 [cité 4 oct 2023]. p. 273-99. (Parasitology Research Monographs). Disponible sur: https://doi.org/10.1007/978-3-642-19382-8_12
12. Martín-Escolano J, Marín C, Rosales MJ, Tsaousis AD, Medina-Carmona E, Martín-Escolano R. An Updated View of the *Trypanosoma cruzi* Life Cycle: Intervention Points for an Effective Treatment. *ACS Infect Dis*. 10 juin 2022;8(6):1107-15.
13. Schwabl P, Imamura H, Van den Broeck F, Costales JA, Manguashca-Sánchez J, Miles MA, et al. Meiotic sex in Chagas disease parasite *Trypanosoma cruzi*. *Nat Commun*. 3 sept 2019;10:3972.
14. Coura JR, Viñas PA. Chagas disease: a new worldwide challenge. *Nature*. juin 2010;465(7301):S6-7.
15. Kemmerling U, Osuna A, Schijman AG, Truyens C. Congenital Transmission of *Trypanosoma cruzi*: A Review About the Interactions Between the Parasite, the Placenta, the Maternal and the Fetal/Neonatal Immune Responses. *Front Microbiol*. 14 août 2019;10:1854.
16. Messenger LA, Gilman RH, Verastegui M, Galdos-Cardenas G, Sanchez G, Valencia E, et al. Toward Improving Early Diagnosis of Congenital Chagas Disease in an Endemic Setting. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*. 7 juill 2017;65(2):268.

17. Pierrotti LC, Carvalho NB, Amorin JP, Pascual J, Kotton CN, López-Vélez R. Chagas Disease Recommendations for Solid-Organ Transplant Recipients and Donors. *Transplantation*. févr 2018;102(2S):S1.
18. Filigheddu MT, Górgolas M, Ramos JM. Enfermedad de Chagas de transmisión oral. *Medicina Clínica*. 9 févr 2017;148(3):125-31.
19. Steindel M, Kramer Pacheco L, Scholl D, Soares M, de Moraes MH, Eger I, et al. Characterization of *Trypanosoma cruzi* isolated from humans, vectors, and animal reservoirs following an outbreak of acute human Chagas disease in Santa Catarina State, Brazil. *Diagnostic Microbiology and Infectious Disease*. 1 janv 2008;60(1):25-32.
20. Messenger LA, Miles MA, Bern C. Between a bug and a hard place: *Trypanosoma cruzi* genetic diversity and the clinical outcomes of Chagas disease. *Expert Rev Anti Infect Ther*. 3 août 2015;13(8):995-1029.
21. Marcili A, Lima L, Cavazzana M, Junqueira ACV, Veludo HH, Silva FMD, et al. A new genotype of *Trypanosoma cruzi* associated with bats evidenced by phylogenetic analyses using SSU rDNA, cytochrome b and Histone H2B genes and genotyping based on ITS1 rDNA. *Parasitology*. mai 2009;136(6):641-55.
22. Zingales B, Miles MA, Campbell DA, Tibayrenc M, Macedo AM, Teixeira MMG, et al. The revised *Trypanosoma cruzi* subspecific nomenclature: Rationale, epidemiological relevance and research applications. *Infection, Genetics and Evolution*. 1 mars 2012;12(2):240-53.
23. Zingales B. *Trypanosoma cruzi* genetic diversity: Something new for something known about Chagas disease manifestations, serodiagnosis and drug sensitivity. *Acta Tropica*. 1 août 2018;184:38-52.
24. Duz ALC, Vieira PM de A, Roatt BM, Aguiar-Soares RDO, Cardoso JM de O, Oliveira FCB de, et al. The TcI and TcII *Trypanosoma cruzi* experimental infections induce distinct immune responses and cardiac fibrosis in dogs. *Mem Inst Oswaldo Cruz*. déc 2014;109:1005-13.
25. Queiroga TBD, Pereira N de S, da Silva DD, Andrade C de M, de Araújo Júnior RF, Brito CR do N, et al. Virulence of *Trypanosoma cruzi* Strains Is Related to the Differential Expression of Innate Immune Receptors in the Heart. *Front Cell Infect Microbiol*. 15 juill 2021;11:696719.
26. Hotez PJ, Dumonteil E, Betancourt Cravioto M, Bottazzi ME, Tapia-Conyer R, Meymandi S, et al. An Unfolding Tragedy of Chagas Disease in North America. *PLoS Negl Trop Dis*. 31 oct 2013;7(10):e2300.
27. Araújo A, Jansen AM, Reinhard K, Ferreira LF. Paleoparasitology of Chagas disease: a review. *Mem Inst Oswaldo Cruz*. juill 2009;104:9-16.
28. Pehrson PO, Wahlgren M, Bengtsson E. Asymptomatic Congenital Chagas' Disease in a 5-Year-Old Child. *Scandinavian Journal of Infectious Diseases*. 1 déc 1981;13(4):307-8.
29. Montalvo-Ocototxle IG, Rojas-Velasco G, Rodríguez-Morales O, Arce-Fonseca M, Baeza-Herrera LA, Arzate-Ramírez A, et al. Chagas Heart Disease: Beyond a Single Complication, from Asymptomatic Disease to Heart Failure. *J Clin Med*. 7 déc 2022;11(24):7262.
30. Gómez-Ochoa SA, Rojas LZ, Echeverría LE, Muka T, Franco OH. Global, Regional, and National Trends of Chagas Disease from 1990 to 2019: Comprehensive Analysis of the Global Burden of Disease Study. *Glob Heart*. 17(1):59.
31. Dias JCP, Ramos Jr. AN, Gontijo ED, Luquetti A, Shikanai-Yasuda MA, Coura JR, et al. 2nd Brazilian Consensus on Chagas Disease, 2015. *Rev Soc Bras Med Trop*. déc 2016;49:03-60.
32. Avaria A, Ventura-Garcia L, Sanmartino M, Van der Laat C. Population movements, borders, and Chagas disease. *Mem Inst Oswaldo Cruz*. 8 juill 2022;117:e210151.

33. Martins-Melo FR, Carneiro M, Ribeiro ALP, Bezerra JMT, Werneck GL. Burden of Chagas disease in Brazil, 1990–2016: findings from the Global Burden of Disease Study 2016. *International Journal for Parasitology*. 1 mars 2019;49(3):301-10.
34. Irish A, Whitman JD, Clark EH, Marcus R, Bern C. Updated Estimates and Mapping for Prevalence of Chagas Disease among Adults, United States. *Emerg Infect Dis*. juill 2022;28(7):1313-20.
35. Lynn MK, Bossak BH, Sandifer PA, Watson A, Nolan MS. Contemporary autochthonous human Chagas disease in the USA. *Acta Tropica*. 1 mai 2020;205:105361.
36. Antinori S, Galimberti L, Bianco R, Grande R, Galli M, Corbellino M. Chagas disease in Europe: A review for the internist in the globalized world. *European Journal of Internal Medicine*. 1 sept 2017;43:6-15.
37. Guerri-Guttenberg RA, Grana DR, Ambrosio G, Milei J. Chagas cardiomyopathy: Europe is not spared!†. *European Heart Journal*. 1 nov 2008;29(21):2587-91.
38. Iglesias Rodríguez IM, Miura S, Maeda T, Imai K, Smith C, Vasquez Velasquez C, et al. Analysis of the Chagas disease situation in Japan: a cross sectional study and cost-effectiveness analysis of a Chagas disease screening program. *Lancet Reg Health West Pac*. 5 sept 2022;31:100574.
39. Lee BY, Bacon KM, Bottazzi ME, Hotez PJ. Global economic burden of Chagas disease: a computational simulation model. *Lancet Infect Dis*. avr 2013;13(4):342-8.
40. Pérez-Molina JA, Molina I. Chagas disease. *The Lancet*. 6 janv 2018;391(10115):82-94.
41. Rassi A, Marin-Neto JA. Chagas disease. *The Lancet*. 17 avr 2010;375(9723):1388-402.
42. Rossi MA, Ramos SG, Bestetti RB. Chagas' heart disease: clinical-pathological correlation. *Front Biosci*. 1 janv 2003;8:e94-109.
43. Echavarría NG, Echeverría LE, Stewart M, Gallego C, Saldarriaga C. Chagas Disease: Chronic Chagas Cardiomyopathy. *Current Problems in Cardiology*. 1 mars 2021;46(3):100507.
44. Medina-Rincón GJ, Gallo-Bernal S, Jiménez PA, Cruz-Saavedra L, Ramírez JD, Rodríguez MJ, et al. Molecular and Clinical Aspects of Chronic Manifestations in Chagas Disease: A State-of-the-Art Review. *Pathogens*. 16 nov 2021;10(11):1493.
45. Nunes MCP, Beaton A, Acquatella H, Bern C, Bolger AF, Echeverría LE, et al. Chagas Cardiomyopathy: An Update of Current Clinical Knowledge and Management: A Scientific Statement From the American Heart Association. *Circulation*. 18 sept 2018;138(12):e169-209.
46. Nunes MCP, Dones W, Morillo CA, Encina JJ, Ribeiro AL. Chagas Disease: An Overview of Clinical and Epidemiological Aspects. *Journal of the American College of Cardiology*. 27 août 2013;62(9):767-76.
47. CVPM SW RVT. *Today's Veterinary Practice*. 2020 [cité 4 oct 2023]. Chagas Disease in Dogs: Transmission, Diagnosis, Treatment, and Prevention. Disponible sur: <https://todaysveterinarypractice.com/parasitology/chagas-disease-dogs/>
48. Freitas HFG, Chizzola PR, Paes ÂT, Lima ACP, Mansur AJ. Risk stratification in a Brazilian hospital-based cohort of 1220 outpatients with heart failure: role of Chagas' heart disease. *International Journal of Cardiology*. 10 juill 2005;102(2):239-47.
49. Matsuda NM, Miller SM, Evora PRB. The Chronic Gastrointestinal Manifestations of Chagas Disease. *Clinics (Sao Paulo)*. déc 2009;64(12):1219-24.
50. Norman FF, López-Vélez R. Chagas disease: comments on the 2018 PAHO Guidelines for diagnosis and management. *Journal of Travel Medicine*. 14 oct 2019;26(7):taz060.
51. Alarcón A, Morgan M, Montgomery SP, Scavo L, Wong ECC, Hahn A, et al. Diagnosis and Treatment of Congenital Chagas Disease in a Premature Infant. *J Pediatric Infect Dis Soc*. déc 2016;5(4):e28-31.

52. Candia-Puma MA, Machaca-Luque LY, Roque-Pumahuanca BM, Galdino AS, Giunchetti RC, Coelho EAF, et al. Accuracy of Diagnostic Tests for the Detection of Chagas Disease: A Systematic Review and Meta-Analysis. *Diagnostics (Basel)*. 10 nov 2022;12(11):2752.
53. Schijman AG, Bisio M, Orellana L, Sued M, Duffy T, Mejia Jaramillo AM, et al. International Study to Evaluate PCR Methods for Detection of *Trypanosoma cruzi* DNA in Blood Samples from Chagas Disease Patients. *PLoS Negl Trop Dis*. 11 janv 2011;5(1):e931.
54. Rojas LZ, Glisic M, Pletsch-Borba L, Echeverría LE, Bramer WM, Bano A, et al. Electrocardiographic abnormalities in Chagas disease in the general population: A systematic review and meta-analysis. *PLoS Negl Trop Dis*. 13 juin 2018;12(6):e0006567.
55. Traina MI, Hernandez S, Sanchez DR, Dufani J, Salih M, Abuhamidah AM, et al. Prevalence of Chagas Disease in a U.S. Population of Latin American Immigrants with Conduction Abnormalities on Electrocardiogram. *PLoS Negl Trop Dis*. 5 janv 2017;11(1):e0005244.
56. Santos É, Menezes Falcão L. Chagas cardiomyopathy and heart failure: From epidemiology to treatment. *Revista Portuguesa de Cardiologia*. 1 mai 2020;39(5):279-89.
57. Ferreira LRP, Ferreira FM, Nakaya HI, Deng X, Cândido D da S, de Oliveira LC, et al. Blood Gene Signatures of Chagas Cardiomyopathy With or Without Ventricular Dysfunction. *J Infect Dis*. 1 févr 2017;215(3):387-95.
58. Orozco VH, Navarro JEV, Aguirre CS, Ocampo CMI, Sandoval CID, Gallegos CAO, et al. Digestive Disorders in Chagas Disease: Megaesophagus and Chagasic Megacolon. In: *Chagas Disease - From Cellular and Molecular Aspects of Trypanosoma cruzi-Host Interactions to the Clinical Intervention* [Internet]. IntechOpen; 2022 [cité 4 oct 2023]. Disponible sur: <https://www.intechopen.com/chapters/80917>
59. Miranda-Arboleda AF, Zaidel EJ, Marcus R, Pinazo MJ, Echeverría LE, Saldarriaga C, et al. Roadblocks in Chagas disease care in endemic and nonendemic countries: Argentina, Colombia, Spain, and the United States. The NET-Heart project. *PLoS Negl Trop Dis*. 30 déc 2021;15(12):e0009954.
60. Cao Z, Jia Y, Zhu B. BNP and NT-proBNP as Diagnostic Biomarkers for Cardiac Dysfunction in Both Clinical and Forensic Medicine. *Int J Mol Sci*. 12 avr 2019;20(8):1820.
61. Ribeiro ALP, Reis AM dos, Barros MVL, Sousa MR de, Rocha ALL, Perez AA, et al. Brain natriuretic peptide and left ventricular dysfunction in Chagas' disease. *The Lancet*. 10 août 2002;360(9331):461-2.
62. Keating S, Deng X, Fernandes F, Cunha-Neto E, Ribeiro A, Adesina B, et al. Inflammatory and Cardiac Biomarkers are Differentially Expressed in Chagas Disease. *Int J Cardiol*. 15 nov 2015;199:451-9.
63. Jacob R, Khan M. Cardiac Biomarkers: What Is and What Can Be. *Indian J Cardiovasc Dis Women WINCARS*. déc 2018;3(4):240-4.
64. Troncoso MF, Ortiz-Quintero J, Garrido-Moreno V, Sanhueza-Olivares F, Guerrero-Moncayo A, Chiong M, et al. VCAM-1 as a predictor biomarker in cardiovascular disease. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*. 1 sept 2021;1867(9):166170.
65. Sousa GR, Gomes JAS, Fares RCG, Damásio MP de S, Chaves AT, Ferreira KS, et al. Plasma Cytokine Expression Is Associated with Cardiac Morbidity in Chagas Disease. *PLoS One*. 6 mars 2014;9(3):e87082.
66. Fares RCG, Gomes J de AS, Garzoni LR, Waghbi MC, Saraiva RM, Medeiros NI, et al. Matrix Metalloproteinases 2 and 9 Are Differentially Expressed in Patients with Indeterminate and Cardiac Clinical Forms of Chagas Disease. *Infect Immun*. oct 2013;81(10):3600-8.

67. Medeiros NI, Gomes JAS, Fiuza JA, Sousa GR, Almeida EF, Novaes RO, et al. MMP-2 and MMP-9 plasma levels are potential biomarkers for indeterminate and cardiac clinical forms progression in chronic Chagas disease. *Sci Rep.* 2 oct 2019;9:14170.
68. Ferreira LRP, Frade AF, Santos RHB, Teixeira PC, Baron MA, Navarro IC, et al. MicroRNAs miR-1, miR-133a, miR-133b, miR-208a and miR-208b are dysregulated in Chronic Chagas disease Cardiomyopathy. *International Journal of Cardiology.* 20 août 2014;175(3):409-17.
69. Pérez-Molina JA, Crespillo-Andújar C, Bosch-Nicolau P, Molina I. Trypanocidal treatment of Chagas disease. *Enfermedades infecciosas y microbiología clinica (English ed).* 1 nov 2021;39(9):458-70.
70. Echeverria LE, Morillo CA. American Trypanosomiasis (Chagas Disease). *Infectious Disease Clinics of North America.* 1 mars 2019;33(1):119-34.
71. Arrúa EC, Seremeta KP, Bedogni GR, Okulik NB, Salomon CJ. Nanocarriers for effective delivery of benznidazole and nifurtimox in the treatment of chagas disease: A review. *Acta Tropica.* 1 oct 2019;198:105080.
72. Morillo CA, Marin-Neto JA, Avezum A, Sosa-Estani S, Rassi A, Rosas F, et al. Randomized Trial of Benznidazole for Chronic Chagas' Cardiomyopathy. *N Engl J Med.* oct 2015;373(14):1295-306.
73. Hasslocher-Moreno AM, Saraiva RM, Sangenis LHC, Xavier SS, de Sousa AS, Costa AR, et al. Benznidazole decreases the risk of chronic Chagas disease progression and cardiovascular events: A long-term follow up study. *EClinicalMedicine.* 23 déc 2020;31:100694.
74. Issa VS, Amaral AF, Cruz FD, Ferreira SMA, Guimarães GV, Chizzola PR, et al. β -Blocker Therapy and Mortality of Patients With Chagas Cardiomyopathy. *Circulation: Heart Failure.* janv 2010;3(1):82-8.
75. Figueiredo CS, de Melo RMV, Viana TT, de Jesus AGQ, da Silva TC, da Silva VM, et al. Clinical and echocardiographic characteristics after six months of sacubitril/valsartan in Chagas heart disease – A case series. *British Journal of Clinical Pharmacology.* 2022;88(2):429-36.
76. Fiorelli AI, Santos RHB, Oliveira JL, Lourenço-Filho DD, Dias RR, Oliveira AS, et al. Heart Transplantation in 107 Cases of Chagas' Disease. *Transplantation Proceedings.* 1 janv 2011;43(1):220-4.
77. Pinazo MJ, Cañas E, Elizalde JI, García M, Gascón J, Gimeno F, et al. Diagnosis, management and treatment of chronic Chagas' gastrointestinal disease in areas where *Trypanosoma cruzi* infection is not endemic. *Gastroenterología y Hepatología.* mars 2010;33(3):191-200.
78. Epting CL, Coatesa BM, Engmanb DM. Molecular Mechanisms of Host Cell Invasion by *Trypanosoma cruzi*. *Exp Parasitol.* nov 2010;126(3):283-91.
79. Calvet CM, Melo TG, Garzoni LR, Oliveira Jr. FOR, Neto DTS, N. S. L. M, et al. Current understanding of the *Trypanosoma cruzi*-cardiomyocyte interaction. *Front Immunol.* 30 oct 2012;3:327.
80. Ferri G, Edreira MM. All Roads Lead to Cytosol: *Trypanosoma cruzi* Multi-Strategic Approach to Invasion. *Front Cell Infect Microbiol.* 5 mars 2021;11:634793.
81. Rodríguez-Bejarano OH, Avedaño C, Patarroyo MA. Mechanisms Associated with *Trypanosoma cruzi* Host Target Cell Adhesion, Recognition and Internalization. *Life (Basel).* 9 juin 2021;11(6):534.
82. Kawai T, Akira S. The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. *Nat Immunol.* mai 2010;11(5):373-84.
83. Macaluso G, Grippi F, Di Bella S, Blanda V, Gucciardi F, Torina A, et al. A Review on the Immunological Response against *Trypanosoma cruzi*. *Pathogens.* 8 févr 2023;12(2):282.

84. Rodrigues MM, Oliveira AC, Bellio M. The Immune Response to *Trypanosoma cruzi*: Role of Toll-Like Receptors and Perspectives for Vaccine Development. *J Parasitol Res.* 2012;2012:507874.
85. Campos MA, Closel M, Valente EP, Cardoso JE, Akira S, Alvarez-Leite JI, et al. Impaired production of proinflammatory cytokines and host resistance to acute infection with *Trypanosoma cruzi* in mice lacking functional myeloid differentiation factor 88. *J Immunol.* 1 févr 2004;172(3):1711-8.
86. Campos MA, Gazzinelli RT. *Trypanosoma cruzi* and its components as exogenous mediators of inflammation recognized through Toll-like receptors. *Mediators Inflamm.* juin 2004;13(3):139-43.
87. Bafica A, Santiago HC, Goldszmid R, Ropert C, Gazzinelli RT, Sher A. Cutting edge: TLR9 and TLR2 signaling together account for MyD88-dependent control of parasitemia in *Trypanosoma cruzi* infection. *J Immunol.* 15 sept 2006;177(6):3515-9.
88. Aoki MP, Carrera-Silva EA, Cuervo H, Fresno M, Gironès N, Gea S. Nonimmune Cells Contribute to Crosstalk between Immune Cells and Inflammatory Mediators in the Innate Response to *Trypanosoma cruzi* Infection. *J Parasitol Res.* 2012;2012:737324.
89. Li R, Jia Z, Trush MA. Defining ROS in Biology and Medicine. *React Oxyg Species (Apex).* 2016;1(1):9-21.
90. Andrieux P, Chevillard C, Cunha-Neto E, Nunes JPS. Mitochondria as a Cellular Hub in Infection and Inflammation. *Int J Mol Sci.* 20 oct 2021;22(21):11338.
91. Maldonado E, Rojas DA, Morales S, Miralles V, Solari A. Dual and Opposite Roles of Reactive Oxygen Species (ROS) in Chagas Disease: Beneficial on the Pathogen and Harmful on the Host. *Oxid Med Cell Longev.* 10 déc 2020;2020:8867701.
92. Hölscher C, Köhler G, Müller U, Mossmann H, Schaub GA, Brombacher F. Defective Nitric Oxide Effector Functions Lead to Extreme Susceptibility of *Trypanosoma cruzi*-Infected Mice Deficient in Gamma Interferon Receptor or Inducible Nitric Oxide Synthase. *Infect Immun.* mars 1998;66(3):1208-15.
93. Chaplin DD. Overview of the Immune Response. *J Allergy Clin Immunol.* févr 2010;125(2 Suppl 2):S3-23.
94. Tai Y, Wang Q, Korner H, Zhang L, Wei W. Molecular Mechanisms of T Cells Activation by Dendritic Cells in Autoimmune Diseases. *Front Pharmacol.* 26 juin 2018;9:642.
95. Sigal LJ. Activation of CD8 T Lymphocytes during Viral Infections. *Encyclopedia of Immunobiology.* 2016;286-90.
96. Acevedo GR, Girard MC, Gómez KA. The Unsolved Jigsaw Puzzle of the Immune Response in Chagas Disease. *Front Immunol.* 2018;9:1929.
97. Megha KB, Mohanan PV. Role of immunoglobulin and antibodies in disease management. *International Journal of Biological Macromolecules.* 1 févr 2021;169:28-38.
98. Rodriguez AM, Santoro F, Afchain D, Bazin H, Capron A. *Trypanosoma cruzi* infection in B-cell-deficient rats. *Infect Immun.* févr 1981;31(2):524-9.
99. Brodskyn C, Silva A, Takehara H, Mota I. IgG subclasses responsible for immune clearance in mice infected with *Trypanosoma cruzi*. *Immunology & Cell Biology.* 1989;67(6):343-8.
100. Rottenberg ME, Bakhiet M, Olsson T, Kristensson K, Mak T, Wigzell H, et al. Differential susceptibilities of mice genomically deleted of CD4 and CD8 to infections with *Trypanosoma cruzi* or *Trypanosoma brucei*. *Infect Immun.* déc 1993;61(12):5129-33.
101. Rodrigues MM, Ribeirão M, Boscardin SB. CD4 Th1 but not Th2 clones efficiently activate macrophages to eliminate *Trypanosoma cruzi* through a nitric oxide dependent mechanism. *Immunology Letters.* 3 juill 2000;73(1):43-50.

102. Hunter CA, Ellis-Neyes LA, Slifer T, Kanaly S, Grünig G, Fort M, et al. IL-10 is required to prevent immune hyperactivity during infection with *Trypanosoma cruzi*. *J Immunol*. 1 avr 1997;158(7):3311-6.
103. Dutra WO, Menezes CAS, Magalhães LMD, Gollob KJ. Immunoregulatory networks in human Chagas disease. *Parasite Immunol*. août 2014;36(8):377-87.
104. Miyazaki Y, Hamano S, Wang S, Shimanoe Y, Iwakura Y, Yoshida H. IL-17 is necessary for host protection against acute-phase *Trypanosoma cruzi* infection. *J Immunol*. 15 juill 2010;185(2):1150-7.
105. Cai CW, Blase JR, Zhang X, Eickhoff CS, Hoft DF. Th17 Cells Are More Protective Than Th1 Cells Against the Intracellular Parasite *Trypanosoma cruzi*. *PLoS Pathog*. 3 oct 2016;12(10):e1005902.
106. Dotiwala F, Mulik S, Polidoro RB, Ansara JA, Burleigh BA, Walch M, et al. Killer lymphocytes use granulysin, perforin and granzymes to kill intracellular parasites. *Nat Med*. févr 2016;22(2):210-6.
107. Tarleton RL, Koller BH, Latour A, Postan M. Susceptibility of β 2-microglobulin-deficient mice to *Trypanosoma cruzi* infection. *Nature*. mars 1992;356(6367):338-40.
108. Kumar S, Tarleton RL. The relative contribution of antibody production and CD8+ T cell function to immune control of *Trypanosoma cruzi*. *Parasite Immunology*. 1998;20(5):207-16.
109. Gorosito Serrán M, Tosello Boari J, Fiocca Vernengo F, Beccaríá CG, Ramello MC, Bermejo DA, et al. Unconventional Pro-inflammatory CD4+ T Cell Response in B Cell-Deficient Mice Infected with *Trypanosoma cruzi*. *Front Immunol*. 21 nov 2017;8:1548.
110. Cardillo F, Postol E, Nihei J, Aroeira LS, Nomizo A, Mengel J. B cells modulate T cells so as to favour T helper type 1 and CD8+ T-cell responses in the acute phase of *Trypanosoma cruzi* infection. *Immunology*. déc 2007;122(4):584-95.
111. Bryan MA, Guyach SE, Norris KA. Specific Humoral Immunity versus Polyclonal B Cell Activation in *Trypanosoma cruzi* Infection of Susceptible and Resistant Mice. *PLoS Negl Trop Dis*. 6 juill 2010;4(7):e733.
112. Tarleton RL, Kuhn RE. Measurement of parasite-specific immune responses in vitro: evidence for suppression of the antibody response to *Trypanosoma cruzi*. *European Journal of Immunology*. 1985;15(8):845-50.
113. Zuñiga E, Motran C, Montes CL, Diaz FL, Bocco JL, Gruppi A. *Trypanosoma cruzi*-induced immunosuppression: B cells undergo spontaneous apoptosis and lipopolysaccharide (LPS) arrests their proliferation during acute infection. *Clin Exp Immunol*. mars 2000;119(3):507-15.
114. Müller U, Schaub GA, Mossmann H, Köhler G, Carsetti R, Hölscher C. Immunosuppression in Experimental Chagas Disease Is Mediated by an Alteration of Bone Marrow Stromal Cell Function During the Acute Phase of Infection. *Front Immunol*. 10 déc 2018;9:2794.
115. Minoprio P, Itohara S, Heusser C, Tonegawa S, Coutinho A. Immunobiology of Murine T *Cruzi* Infection: The Predominance of Parasite-nonspecific Responses and the Activation of TCRIT Cells. *Immunological Reviews*. 1989;112(1):183-207.
116. Minoprio P, Burlen O, Pereira P, Guilbert B, Andrade L, Hontebeyrie-Joskowicz M, et al. Most B Cells in Acute *Trypanosoma cruzi* Infection Lack Parasite Specificity. *Scandinavian Journal of Immunology*. 1988;28(5):553-61.
117. Bilate AMB, Previato JO, Mendonça-Previato L, Peçanha LMT. Glycoinositolphospholipids from *Trypanosoma cruzi* induce B cell hyper-responsiveness in vivo. *Glycoconj J*. 1 oct 2000;17(10):727-34.
118. Nothelfer K, Sansonetti PJ, Phalipon A. Pathogen manipulation of B cells: the best defence is a good offence. *Nat Rev Microbiol*. mars 2015;13(3):173-84.

119. Piacenza L, Peluffo G, Alvarez MN, Martínez A, Radi R. Trypanosoma cruzi Antioxidant Enzymes As Virulence Factors in Chagas Disease. *Antioxid Redox Signal*. 1 sept 2013;19(7):723-34.
120. Gonçalves RLS, Barreto RFSM, Polycarpo CR, Gadelha FR, Castro SL, Oliveira MF. A comparative assessment of mitochondrial function in epimastigotes and bloodstream trypomastigotes of Trypanosoma cruzi. *J Bioenerg Biomembr*. 1 déc 2011;43(6):651-61.
121. Paiva CN, Medei E, Bozza MT. ROS and Trypanosoma cruzi: Fuel to infection, poison to the heart. *PLoS Pathog*. 19 avr 2018;14(4):e1006928.
122. Paiva CN, Feijó DF, Dutra FF, Carneiro VC, Freitas GB, Alves LS, et al. Oxidative stress fuels Trypanosoma cruzi infection in mice. *J Clin Invest*. 2 juill 2012;122(7):2531-42.
123. Wan X, Gupta S, Zago MP, Davidson MM, Dousset P, Amoroso A, et al. Defects of mtDNA Replication Impaired Mitochondrial Biogenesis During Trypanosoma cruzi Infection in Human Cardiomyocytes and Chagasic Patients: The Role of Nrf1/2 and Antioxidant Response. *J Am Heart Assoc*. 19 déc 2012;1(6):e003855.
124. Gupta S, Bhatia V, Wen J jun, Wu Y, Huang MH, Garg NJ. Trypanosoma cruzi infection disturbs mitochondrial membrane potential and ROS production rate in cardiomyocytes. *Free Radic Biol Med*. 15 nov 2009;47(10):1414-21.
125. Doyle PS, Zhou YM, Hsieh I, Greenbaum DC, McKerrow JH, Engel JC. The Trypanosoma cruzi Protease Cruzain Mediates Immune Evasion. *PLoS Pathog*. 1 sept 2011;7(9):e1002139.
126. Alcaide P, Fresno M. The Trypanosoma cruzi membrane mucin AgC10 inhibits T cell activation and IL-2 transcription through L-selectin. *International Immunology*. 1 oct 2004;16(10):1365-75.
127. Ruiz Díaz P, Mucci J, Meira MA, Bogliotti Y, Musikant D, Leguizamón MS, et al. Trypanosoma cruzi trans-Sialidase Prevents Elicitation of Th1 Cell Response via Interleukin 10 and Downregulates Th1 Effector Cells. *Infect Immun*. mai 2015;83(5):2099-108.
128. Somoza M, Bertelli A, Pratto CA, Verdun RE, Campetella O, Mucci J. Trypanosoma cruzi Induces B Cells That Regulate the CD4+ T Cell Response. *Front Cell Infect Microbiol*. 5 janv 2022;11:789373.
129. Magalhães LMD, Viana A, Chiari E, Galvão LMC, Gollob KJ, Dutra WO. Differential Activation of Human Monocytes and Lymphocytes by Distinct Strains of Trypanosoma cruzi. *PLoS Negl Trop Dis*. 6 juill 2015;9(7):e0003816.
130. Cristovão-Silva AC, Brelaz-de-Castro MCA, Hernandez MZ, Pereira VRA. Chagas disease: Immunology of the disease at a glance. *Cytokine & Growth Factor Reviews*. 1 déc 2021;62:15-22.
131. Passos LSA, Magalhães LMD, Soares RP, Marques AF, Alves MLR, Giunchetti RC, et al. Activation of Human CD11b+ B1 B-Cells by Trypanosoma cruzi-Derived Proteins Is Associated With Protective Immune Response in Human Chagas Disease. *Front Immunol*. 4 janv 2019;9:3015.
132. Cunha-Neto E, Chevillard C. Chagas Disease Cardiomyopathy: Immunopathology and Genetics. *Mediators Inflamm*. 2014;2014:683230.
133. Magalhães LMD, Villani FNA, Nunes M do CP, Gollob KJ, Rocha MOC, Dutra WO. High Interleukin 17 Expression Is Correlated With Better Cardiac Function in Human Chagas Disease. *J Infect Dis*. 15 févr 2013;207(4):661-5.
134. Guedes PM da M, Gutierrez FRS, Maia FL, Milanezi CM, Silva GK, Pavanelli WR, et al. IL-17 Produced during Trypanosoma cruzi Infection Plays a Central Role in Regulating Parasite-Induced Myocarditis. *PLoS Negl Trop Dis*. 16 févr 2010;4(2):e604.
135. Pérez-Fuentes R, Guégan JF, Barnabé C, López-Colombo A, Salgado-Rosas H, Torres-Rasgado E, et al. Severity of chronic Chagas disease is associated with cytokine/antioxidant imbalance in chronically infected individuals. *International Journal for Parasitology*. 1 mars 2003;33(3):293-9.

136. Vyatkina G, Bhatia V, Gerstner A, Papaconstantinou J, Garg N. Impaired mitochondrial respiratory chain and bioenergetics during chagasic cardiomyopathy development. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*. 28 juin 2004;1689(2):162-73.
137. Wen J jun, Yachelini PC, Sembaj A, Manzur RE, Garg NJ. Increased oxidative stress is correlated with mitochondrial dysfunction in chagasic patients. *Free Radical Biology and Medicine*. 15 juill 2006;41(2):270-6.
138. Silverio JC, Pereira IR, Cipitelli M da C, Vinagre NF, Rodrigues MM, Gazzinelli RT, et al. CD8+ T-Cells Expressing Interferon Gamma or Perforin Play Antagonistic Roles in Heart Injury in Experimental *Trypanosoma Cruzi*-Elicited Cardiomyopathy. *PLoS Pathog*. 19 avr 2012;8(4):e1002645.
139. Laucella SA, Postan M, Martin D, Fralish BH, Albareda MC, Alvarez MG, et al. Frequency of Interferon- γ -Producing T Cells Specific for *Trypanosoma cruzi* Inversely Correlates with Disease Severity in Chronic Human Chagas Disease. *The Journal of Infectious Diseases*. 1 mars 2004;189(5):909-18.
140. Lasso P, Mateus J, Pavia P, Rosas F, Roa N, Thomas MC, et al. Inhibitory Receptor Expression on CD8+ T Cells Is Linked to Functional Responses against *Trypanosoma cruzi* Antigens in Chronic Chagasic Patients. *J Immunol*. 15 oct 2015;195(8):3748-58.
141. Bonney KM, Engman DM. Autoimmune Pathogenesis of Chagas Heart Disease. *Am J Pathol*. juin 2015;185(6):1537-47.
142. Cunha-Neto E, Coelho V, Guilherme L, Fiorelli A, Stolf N, Kalil J. Autoimmunity in Chagas' disease. Identification of cardiac myosin-B13 *Trypanosoma cruzi* protein crossreactive T cell clones in heart lesions of a chronic Chagas' cardiomyopathy patient. *J Clin Invest*. 15 oct 1996;98(8):1709-12.
143. Giordanengo L, Fretes R, Díaz H, Cano R, Bacile A, Vottero-Cima E, et al. Cruzipain induces autoimmune response against skeletal muscle and tissue damage in mice. *Muscle & Nerve*. 2000;23(9):1407-13.
144. Vermelho AB, de Nazareth Leal de Meirelles M, Pereira MC, Pohlentz G, Barreto-Bergter E. Heart muscle cells share common neutral glycosphingolipids with *Trypanosoma cruzi*. *Acta Tropica*. 15 avr 1997;64(3):131-43.
145. Goin JC, Borda E, Segovia A, Sterin-Borda L. Distribution of Antibodies against β -Adrenoceptors in the Course of Human *Trypanosoma cruzi* Infection. *Proceedings of the Society for Experimental Biology and Medicine*. 1 juin 1991;197(2):186-92.
146. Goin JC, Borda E, Perez Leiros C, Storino R, Sterin-Borda L. Identification of antibodies with muscarinic cholinergic activity in human Chagas' disease: pathological implications. *Journal of the Autonomic Nervous System*. 1 avr 1994;47(1):45-52.
147. De Bona E, Lidani KCF, Bavia L, Omidian Z, Gremski LH, Sandri TL, et al. Autoimmunity in Chronic Chagas Disease: A Road of Multiple Pathways to Cardiomyopathy? *Front Immunol*. 6 août 2018;9:1842.
148. Ricci AD, Bracco L, Salas-Sarduy E, Ramsey JM, Nolan MS, Lynn MK, et al. The *Trypanosoma cruzi* Antigen and Epitope Atlas: antibody specificities in Chagas disease patients across the Americas. *Nat Commun*. 3 avr 2023;14:1850.
149. Voineagu I, Wang X, Johnston P, Lowe JK, Tian Y, Horvath S, et al. Transcriptomic Analysis of Autistic Brain Reveals Convergent Molecular Pathology. *Nature*. 25 mai 2011;474(7351):380-4.
150. Saxena A, Mathur N, Tiwari P, Mathur SK. Whole transcriptome RNA-seq reveals key regulatory factors involved in type 2 diabetes pathology in peripheral fat of Asian Indians. *Sci Rep*. 20 mai 2021;11(1):10632.

151. Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science*. 20 oct 1995;270(5235):467-70.
152. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet*. janv 2009;10(1):57-63.
153. Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, et al. mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods*. mai 2009;6(5):377-82.
154. Marx V. Method of the Year: spatially resolved transcriptomics. *Nat Methods*. janv 2021;18(1):9-14.
155. Ciccocanti F, Antonioli M, Sacchi A, Notari S, Farina A, Beccacece A, et al. Proteomic analysis identifies a signature of disease severity in the plasma of COVID-19 pneumonia patients associated to neutrophil, platelet and complement activation. *Clin Proteomics*. 8 nov 2022;19:38.
156. Colak D, Alaiya AA, Kaya N, Muiya NP, AlHarazi O, Shinwari Z, et al. Integrated Left Ventricular Global Transcriptome and Proteome Profiling in Human End-Stage Dilated Cardiomyopathy. *PLoS One*. 6 oct 2016;11(10):e0162669.
157. Shah SH, Bain JR, Muehlbauer MJ, Stevens RD, Crosslin DR, Haynes C, et al. Association of a Peripheral Blood Metabolic Profile With Coronary Artery Disease and Risk of Subsequent Cardiovascular Events. *Circulation: Cardiovascular Genetics*. avr 2010;3(2):207-14.
158. Wang X, Pankratz VS, Fredericksen Z, Tarrell R, Karaus M, McGuffog L, et al. Common variants associated with breast cancer in genome-wide association studies are modifiers of breast cancer risk in BRCA1 and BRCA2 mutation carriers. *Hum Mol Genet*. 15 juill 2010;19(14):2886-97.
159. Martyn GE, Wienert B, Yang L, Shah M, Norton LJ, Burdach J, et al. Natural regulatory mutations elevate the fetal globin gene via disruption of BCL11A or ZBTB7A binding. *Nat Genet*. avr 2018;50(4):498-503.
160. De Carvalho DD, Sharma S, You JS, Su SF, Taberlay PC, Kelly TK, et al. DNA methylation screening identifies driver epigenetic events of cancer cell survival. *Cancer Cell*. 25 mai 2012;21(5):655-67.
161. Krolevets M, Cate V ten, Prochaska JH, Schulz A, Rapp S, Tenzer S, et al. DNA methylation and cardiovascular disease in humans: a systematic review and database of known CpG methylation sites. *Clin Epigenetics*. 30 mars 2023;15:56.
162. Zhao Z, Shilatifard A. Epigenetic modifications of histones in cancer. *Genome Biol*. 20 nov 2019;20:245.
163. Kraus L, Beavens B. The Current Therapeutic Role of Chromatin Remodeling for the Prognosis and Treatment of Heart Failure. *Biomedicines*. 16 févr 2023;11(2):579.
164. Gurtan AM, Sharp PA. The role of miRNAs in regulating gene expression networks. *J Mol Biol*. 9 oct 2013;425(19):3582-600.
165. Colpaert RMW, Calore M. MicroRNAs in Cardiac Diseases. *Cells*. 18 juill 2019;8(7):737.
166. Statello L, Guo CJ, Chen LL, Huarte M. Gene regulation by long non-coding RNAs and its biological functions. *Nat Rev Mol Cell Biol*. 2021;22(2):96-118.
167. Li F, Li X, Qiao L, Liu W, Xu C, Wang X. MALAT1 regulates miR-34a expression in melanoma cells. *Cell Death Dis*. 17 mai 2019;10(6):389.
168. Fang Y, Xu Y, Wang R, Hu L, Guo D, Xue F, et al. Recent advances on the roles of LncRNAs in cardiovascular disease. *J Cell Mol Med*. nov 2020;24(21):12246-57.
169. Qu X, Du Y, Shu Y, Gao M, Sun F, Luo S, et al. MIAT Is a Pro-fibrotic Long Non-coding RNA Governing Cardiac Fibrosis in Post-infarct Myocardium. *Sci Rep*. 15 févr 2017;7:42657.
170. Imker HJ. 25 Years of Molecular Biology Databases: A Study of Proliferation, Impact, and Maintenance. *Frontiers in Research Metrics and Analytics [Internet]*. 2018 [cité 4 oct 2023];3. Disponible sur: <https://www.frontiersin.org/articles/10.3389/frma.2018.00018>

171. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene Ontology: tool for the unification of biology. *Nat Genet.* mai 2000;25(1):25-9.
172. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* 1 janv 2000;28(1):27-30.
173. Hamosh A, Scott AF, Amberger JS, Bocchini CA, McKusick VA. Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. *Nucleic Acids Res.* 1 janv 2005;33(Database Issue):D514-7.
174. Piñero J, Queralt-Rosinach N, Bravo À, Deu-Pons J, Bauer-Mehren A, Baron M, et al. DisGeNET: a discovery platform for the dynamical exploration of human diseases and their genes. *Database (Oxford).* 15 avr 2015;2015:bav028.
175. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, et al. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* 1 janv 2001;29(1):308-11.
176. Huang Z, Shi J, Gao Y, Cui C, Zhang S, Li J, et al. HMDD v3.0: a database for experimentally supported human microRNA–disease associations. *Nucleic Acids Res.* 8 janv 2019;47(Database issue):D1013-7.
177. Chen G, Wang Z, Wang D, Qiu C, Liu M, Chen X, et al. LncRNADisease: a database for long-non-coding RNA-associated diseases. *Nucleic Acids Res.* janv 2013;41(Database issue):D983-6.
178. Xing J, Zhai R, Wang C, Liu H, Zeng J, Zhou D, et al. DiseaseMeth version 3.0: a major expansion and update of the human disease methylation database. *Nucleic Acids Res.* 18 nov 2021;50(D1):D1208-15.
179. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 8 janv 2019;47(Database issue):D607-13.
180. Chatr-aryamontri A, Ceol A, Palazzi LM, Nardelli G, Schneider MV, Castagnoli L, et al. MINT: the Molecular INTeraction database. *Nucleic Acids Res.* janv 2007;35(Database issue):D572-4.
181. Oughtred R, Rust J, Chang C, Breitkreutz B, Stark C, Willems A, et al. The BioGRID database: A comprehensive biomedical resource of curated protein, genetic, and chemical interactions. *Protein Sci.* janv 2021;30(1):187-200.
182. del Toro N, Shrivastava A, Ragueneau E, Meldal B, Combe C, Barrera E, et al. The IntAct database: efficient access to fine-grained molecular interaction data. *Nucleic Acids Res.* 11 nov 2021;50(D1):D648-53.
183. Fang L, Li Y, Ma L, Xu Q, Tan F, Chen G. GRNdb: decoding the gene regulatory networks in diverse human and mouse conditions. *Nucleic Acids Res.* 5 nov 2020;49(D1):D97-103.
184. Yang S, Kim CY, Hwang S, Kim E, Kim H, Shim H, et al. COEXPEDIA: exploring biomedical hypotheses via co-expressions associated with medical subject headings (MeSH). *Nucleic Acids Res.* 4 janv 2017;45(Database issue):D389-96.
185. Hammal F, de Langen P, Bergon A, Lopez F, Ballester B. ReMap 2022: a database of Human, Mouse, Drosophila and Arabidopsis regulatory regions from an integrative analysis of DNA-binding sequencing experiments. *Nucleic Acids Res.* 9 nov 2021;50(D1):D316-25.
186. Huang HY, Lin YCD, Cui S, Huang Y, Tang Y, Xu J, et al. miRTarBase update 2022: an informative resource for experimentally validated miRNA–target interactions. *Nucleic Acids Res.* 30 nov 2021;50(D1):D222-30.
187. Karagkouni D, Paraskevopoulou MD, Chatzopoulos S, Vlachos IS, Tastsoglou S, Kanellos I, et al. DIANA-TarBase v8: a decade-long collection of experimentally supported miRNA–gene interactions. *Nucleic Acids Res.* 4 janv 2018;46(Database issue):D239-45.
188. Xiao F, Zuo Z, Cai G, Kang S, Gao X, Li T. miRecords: an integrated resource for microRNA–target interactions. *Nucleic Acids Res.* janv 2009;37(Database issue):D105-10.

189. Zhao H, Yin X, Xu H, Liu K, Liu W, Wang L, et al. LncTarD 2.0: an updated comprehensive database for experimentally-supported functional lncRNA–target regulations in human diseases. *Nucleic Acids Res.* 2 nov 2022;51(D1):D199-207.
190. Hauschka TS. Sex of Host as a Factor in Chagas' Disease. *The Journal of Parasitology.* 1947;33(5):399-404.
191. Zicker F, Smith PG, Netto JCA, Oliveira RM, Zicker EMS. Physical Activity, Opportunity for Reinfection, and Sibling History of Heart Disease as Risk Factors for Chagas' Cardiopathy. *The American Journal of Tropical Medicine and Hygiene.* 1 nov 1990;43(5):498-505.
192. Wrightsman R, Krassner S, Watson J. Genetic control of responses to *Trypanosoma cruzi* in mice: multiple genes influencing parasitemia and survival. *Infect Immun.* mai 1982;36(2):637-44.
193. Boyer MH, Hoff R, Kipnis TL, Murphy ED, Roths JB. *Trypanosoma cruzi*: Susceptibility in mice carrying mutant gene *Ipr* (lymphoproliferation). *Parasite Immunology.* 1983;5(2):135-42.
194. Trischmann TM. Single locus in BXH-2 mice responsible for inability to control early proliferation of *Trypanosoma cruzi*. *Infect Immun.* déc 1984;46(3):658-62.
195. Layrisse Z, Fernandez MT, Montagnani S, Matos M, Balbas O, Herrera F, et al. HLA-C*03 is a risk factor for cardiomyopathy in chagas disease. *Human Immunology.* 1 sept 2000;61(9):925-9.
196. Nieto A, Beraún Y, Callado M d., Caballero A, Alonso A, González A, et al. HLA haplotypes are associated with differential susceptibility to *Trypanosoma cruzi* infection. *Tissue Antigens.* 2000;55(3):195-8.
197. del Puerto F, Nishizawa JE, Kikuchi M, Roca Y, Avilas C, Gianella A, et al. Protective Human Leucocyte Antigen Haplotype, HLA-DRB1*01-B*14, against Chronic Chagas Disease in Bolivia. *PLoS Negl Trop Dis.* 20 mars 2012;6(3):e1587.
198. Dias FC, Mendes-Junior CT, Silva MC, Tristão FSM, Dellalibera-Joviliano R, Moreau P, et al. Human Leucocyte Antigen-G (HLA-G) and Its Murine Functional Homolog Qa2 in the *Trypanosoma cruzi* Infection. *Mediators Inflamm.* 2015;2015:595829.
199. Ramasawmy R, Cunha-Neto E, Faé KC, Müller NG, Cavalcanti VL, Drigo SA, et al. BAT1, a Putative Anti-Inflammatory Gene, Is Associated with Chronic Chagas Cardiomyopathy. *The Journal of Infectious Diseases.* 15 mai 2006;193(10):1394-9.
200. Ramasawmy R, Cunha-Neto E, Faé KC, Martello FG, Müller NG, Cavalcanti VL, et al. The Monocyte Chemoattractant Protein–1 Gene Polymorphism Is Associated with Cardiomyopathy in Human Chagas Disease. *Clinical Infectious Diseases.* 1 août 2006;43(3):305-11.
201. Frade AF, Pissetti CW, Ianni BM, Saba B, Lin-Wang HT, Nogueira LG, et al. Genetic susceptibility to Chagas disease cardiomyopathy: involvement of several genes of the innate immunity and chemokine-dependent migration pathways. *BMC Infect Dis.* 12 déc 2013;13:587.
202. Criado L, Flórez O, Martín J, González CI. Genetic polymorphisms in TNFA/TNFR2 genes and Chagas disease in a Colombian endemic population. *Cytokine.* 1 mars 2012;57(3):398-401.
203. Sandri TL, Andrade FA, Lidani KCF, Einig E, Boldt ABW, Mordmüller B, et al. Human collectin-11 (COLEC11) and its synergic genetic interaction with MASP2 are associated with the pathophysiology of Chagas Disease. *PLoS Negl Trop Dis.* 17 avr 2019;13(4):e0007324.
204. Calzada J e., Nieto A, Beraún Y, Martín J. Chemokine receptor CCR5 polymorphisms and Chagas' disease cardiomyopathy. *Tissue Antigens.* 2001;58(3):154-8.
205. Sandri TL, Lidani KCF, Andrade FA, Meyer CG, Kremsner PG, de Messias-Reason IJ, et al. Human complement receptor type 1 (CR1) protein levels and genetic variants in chronic Chagas Disease. *Sci Rep.* 11 janv 2018;8(1):526.
206. Dias FC, Medina T da S, Mendes-Junior CT, Dantas RO, Pissetti CW, Rodrigues Junior V, et al. Polymorphic Sites at the Immunoregulatory CTLA-4 Gene Are Associated with Chronic Chagas Disease and Its Clinical Manifestations. *PLoS One.* 24 oct 2013;8(10):e78367.

207. Costa GC, da Costa Rocha MO, Moreira PR, Menezes AS, Silva MR, Gollob KJ, et al. Functional IL-10 Gene Polymorphism Is Associated with Chagas Disease Cardiomyopathy. *The Journal of Infectious Diseases*. 1 févr 2009;199(3):451-4.
208. Grijalva A, Gallo Vaulet L, Agüero RN, Toledano A, Risso MG, Quarroz Braghini J, et al. Interleukin 10 Polymorphisms as Risk Factors for Progression to Chagas Disease Cardiomyopathy: A Case-Control Study and Meta-Analysis. *Front Immunol*. 4 juill 2022;13:946350.
209. Frade-Barros AF, Ianni BM, Cabantous S, Pissetti CW, Saba B, Lin-Wang HT, et al. Polymorphisms in Genes Affecting Interferon- γ Production and Th1 T Cell Differentiation Are Associated With Progression to Chagas Disease Cardiomyopathy. *Front Immunol*. 7 juill 2020;11:1386.
210. Flórez O, Zafra G, Morillo C, Martín J, González CI. Interleukin-1 Gene Cluster Polymorphism in Chagas Disease in a Colombian Case-Control Study. *Human Immunology*. 1 sept 2006;67(9):741-8.
211. Leon Rodriguez DA, Echeverría LE, González CI, Martín J. Investigation of the role of IL17A gene variants in Chagas disease. *Genes Immun*. déc 2015;16(8):536-40.
212. Zafra G, Morillo C, Martín J, González A, González CI. Polymorphism in the 3' UTR of the IL12B gene is associated with Chagas' disease cardiomyopathy. *Microbes and Infection*. 1 juill 2007;9(9):1049-52.
213. Strauss M, Acosta-Herrera M, Alcaraz A, Casares-Marfil D, Bosch-Nicolau P, Lo Presti MS, et al. Association of IL18 genetic polymorphisms with Chagas disease in Latin American populations. *PLoS Negl Trop Dis*. 21 nov 2019;13(11):e0007859.
214. Cruz-Robles D, Chávez-González JP, Cavazos-Quero MM, Pérez-Méndez O, Reyes PA, Vargas-Alarcón G. Association Between IL-1B and IL-1RN Gene Polymorphisms and Chagas' Disease Development Susceptibility. *Immunological Investigations*. 1 janv 2009;38(3-4):231-9.
215. Ramasawmy R, Fae KC, Neto EC, Müller NG, Cavalcanti VL, Ferreira RC, et al. Polymorphisms in the Gene for Lymphotoxin- α Predispose to Chronic Chagas Cardiomyopathy. *The Journal of Infectious Diseases*. 15 déc 2007;196(12):1836-43.
216. Luz PR, Miyazaki MI, Chiminacio Neto N, Padeski MC, Barros ACM, Boldt ABW, et al. Genetically Determined MBL Deficiency Is Associated with Protection against Chronic Cardiomyopathy in Chagas Disease. *PLoS Negl Trop Dis*. 8 janv 2016;10(1):e0004257.
217. Ramasawmy R, Cunha-Neto E, Fae KC, Borba SCP, Teixeira PC, Ferreira SCP, et al. Heterozygosity for the S180L Variant of MAL/TIRAP, a Gene Expressing an Adaptor Protein in the Toll-Like Receptor Pathway, Is Associated with Lower Risk of Developing Chronic Chagas Cardiomyopathy. *The Journal of Infectious Diseases*. 15 juin 2009;199(12):1838-45.
218. Ramasawmy R, Faé KC, Cunha-Neto E, Borba SCP, Ianni B, Mady C, et al. Variants in the promoter region of IKBL/NFKBIL1 gene may mark susceptibility to the development of chronic Chagas' cardiomyopathy among *Trypanosoma cruzi*-infected individuals. *Molecular Immunology*. 1 janv 2008;45(1):283-8.
219. Silva MC, da Silva Medina T, Fuzo CA, Dias FC, Freitas-Castro F, Fukutani KF, et al. Polymorphism in the catalytic subunit of the PI3K γ gene is associated with *Trypanosoma cruzi*-induced chronic chagasic cardiomyopathy. *Infection, Genetics and Evolution*. 1 mars 2021;88:104671.
220. Calzada JE, Beraún Y, González CI, Martín J. Transforming growth factor beta 1 (TGF β 1) gene polymorphisms and Chagas disease susceptibility in Peruvian and Colombian patients. *Cytokine*. 1 mars 2009;45(3):149-53.

221. Rodríguez-Pérez JM, Cruz-Robles D, Hernández-Pacheco G, Pérez-Hernández N, Murguía LE, Granados J, et al. Tumor necrosis factor-alpha promoter polymorphism in Mexican patients with Chagas' disease. *Immunology Letters*. 15 avr 2005;98(1):97-102.
222. Leon Rodriguez DA, Carmona FD, González CI, Martin J. Evaluation of VDR gene polymorphisms in *Trypanosoma cruzi* infection and chronic Chagasic cardiomyopathy. *Sci Rep*. 9 août 2016;6:31263.
223. Frade AF, Teixeira PC, Ianni BM, Pissetti CW, Saba B, Wang LHT, et al. Polymorphism in the Alpha Cardiac Muscle Actin 1 Gene Is Associated to Susceptibility to Chronic Inflammatory Cardiomyopathy. *PLoS One*. 19 déc 2013;8(12):e83446.
224. Luz PR, Boldt ABW, Grisbach C, Kun JFJ, Velavan TP, Messias-Reason IJT. Association of L-Ficolin Levels and FCN2 Genotypes with Chronic Chagas Disease. *PLoS One*. 4 avr 2013;8(4):e60237.
225. Nogueira LG, Santos RHB, Ianni BM, Fiorelli AI, Mairena EC, Benvenuti LA, et al. Myocardial Chemokine Expression and Intensity of Myocarditis in Chagas Cardiomyopathy Are Controlled by Polymorphisms in CXCL9 and CXCL10. *PLoS Negl Trop Dis*. 25 oct 2012;6(10):e1867.
226. Machuca MA, Suárez EU, Echeverría LE, Martín J, González CI. SNP/haplotype associations of CCR2 and CCR5 genes with severity of chagasic cardiomyopathy. *Human Immunology*. 1 déc 2014;75(12):1210-5.
227. Fu KYJ, Zamudio R, Henderson-Frost J, Almuedo A, Steinberg H, Clipman SJ, et al. Association of caspase-1 polymorphisms with Chagas cardiomyopathy among individuals in Santa Cruz, Bolivia. *Rev Soc Bras Med Trop*. 2017;50(4):516-23.
228. Reis PG, Ayo CM, de Mattos LC, Brandão de Mattos C de C, Sakita KM, de Moraes AG, et al. Genetic Polymorphisms of IL17 and Chagas Disease in the South and Southeast of Brazil. *J Immunol Res*. 2017;2017:1017621.
229. Nogueira LG, Frade AF, Ianni BM, Laugier L, Pissetti CW, Cabantous S, et al. Functional IL18 polymorphism and susceptibility to Chronic Chagas Disease. *Cytokine*. 1 mai 2015;73(1):79-83.
230. Clipman SJ, Henderson-Frost J, Fu KY, Bern C, Flores J, Gilman RH. Genetic association study of NLRP1, CARD, and CASP1 inflammasome genes with chronic Chagas cardiomyopathy among *Trypanosoma cruzi* seropositive patients in Bolivia. *PLoS One*. 13 févr 2018;13(2):e0192378.
231. Blasco RL, Strauss M, Velázquez López DA, Tabares S, Sembaj A, Rivarola HW, et al. SCN5A gene variants as potential markers of the progression of chronic chagasic cardiac alterations. *Parasitology International*. 1 févr 2021;80:102213.
232. Deng X, Sabino EC, Cunha-Neto E, Ribeiro AL, Ianni B, Mady C, et al. Genome Wide Association Study (GWAS) of Chagas Cardiomyopathy in *Trypanosoma cruzi* Seropositive Subjects. *PLoS One*. 20 nov 2013;8(11):e79629.
233. Casares-Marfil D, Strauss M, Bosch-Nicolau P, Lo Presti MS, Molina I, Chevillard C, et al. A Genome-Wide Association Study Identifies Novel Susceptibility loci in Chronic Chagas Cardiomyopathy. *Clin Infect Dis*. 4 févr 2021;73(4):672-9.
234. Sabino EC, Franco LAM, Venturini G, Velho Rodrigues M, Marques E, de Oliveira-da Silva LC, et al. Genome-wide association study for Chagas Cardiomyopathy identify a new risk locus on chromosome 18 associated with an immune-related protein and transcriptional signature. *PLoS Negl Trop Dis*. 10 oct 2022;16(10):e0010725.
235. Ouarhache M, Marquet S, Frade AF, Ferreira AM, Ianni B, Almeida RR, et al. Rare Pathogenic Variants in Mitochondrial and Inflammation-Associated Genes May Lead to Inflammatory Cardiomyopathy in Chagas Disease. *J Clin Immunol*. 2021;41(5):1048-63.

236. Ferreira BL, Ferreira ÉR, de Brito MV, Salu BR, Oliva MLV, Mortara RA, et al. BALB/c and C57BL/6 Mice Cytokine Responses to *Trypanosoma cruzi* Infection Are Independent of Parasite Strain Infectivity. *Front Microbiol.* 26 mars 2018;9:553.
237. Dutra WO, Gollob KJ, Pinto-Dias JC, Gazzinelli G, Correa-Oliveira R, Coffman RL, et al. Cytokine mRNA Profile of Peripheral Blood Mononuclear Cells Isolated from Individuals with *Trypanosoma cruzi* Chronic Infection. *Scandinavian Journal of Immunology.* 1997;45(1):74-80.
238. Ferreira RC, Ianni BM, Abel LC, Buck P, Mady C, Kalil J, et al. Increased plasma levels of tumor necrosis factor-alpha in asymptomatic/"indeterminate" and Chagas disease cardiomyopathy patients. *Mem Inst Oswaldo Cruz.* avr 2003;98:407-12.
239. Talvani A, Rocha MOC, Barcelos LS, Gomes YM, Ribeiro AL, Teixeira MM. Elevated Concentrations of CCL2 and Tumor Necrosis Factor— α in Chagasic Cardiomyopathy. *Clinical Infectious Diseases.* 1 avr 2004;38(7):943-50.
240. Cunha-Neto E, Dzau VJ, Allen PD, Stamatou D, Benvenuti L, Higuchi ML, et al. Cardiac Gene Expression Profiling Provides Evidence for Cytokinopathy as a Molecular Mechanism in Chagas' Disease Cardiomyopathy. *Am J Pathol.* août 2005;167(2):305-13.
241. Teixeira PC, Santos RHB, Fiorelli AI, Bilate AMB, Benvenuti LA, Stolf NA, et al. Selective Decrease of Components of the Creatine Kinase System and ATP Synthase Complex in Chronic Chagas Disease Cardiomyopathy. *PLoS Negl Trop Dis.* 28 juin 2011;5(6):e1205.
242. Nogueira LG, Santos RHB, Fiorelli AI, Mairena EC, Benvenuti LA, Bocchi EA, et al. Myocardial Gene Expression of T-bet, GATA-3, Ror- γ t, FoxP3, and Hallmark Cytokines in Chronic Chagas Disease Cardiomyopathy: An Essentially Unopposed TH1-Type Response. *Mediators Inflamm.* 2014;2014:914326.
243. Gómez-Olarte S, Bolaños NI, Echeverry M, Rodríguez AN, Cuéllar A, Puerta CJ, et al. Intermediate Monocytes and Cytokine Production Associated With Severe Forms of Chagas Disease. *Front Immunol.* 19 juill 2019;10:1671.
244. Ribeiro AL, Nunes MP, Teixeira MM, Rocha MOC. Diagnosis and management of Chagas disease and cardiomyopathy. *Nat Rev Cardiol.* oct 2012;9(10):576-89.
245. Gomes JAS, Bahia-Oliveira LMG, Rocha MOC, Busek SCU, Teixeira MM, Silva JS, et al. Type 1 Chemokine Receptor Expression in Chagas' Disease Correlates with Morbidity in Cardiac Patients. *Infect Immun.* déc 2005;73(12):7960-6.
246. Cutrullis RA, Petray PB, Schapachnik E, Sánchez R, Postan M, González MN, et al. Elevated Serum Levels of Macrophage Migration Inhibitory Factor Are Associated with Progressive Chronic Cardiomyopathy in Patients with Chagas Disease. *PLoS One.* 22 févr 2013;8(2):e57181.
247. Curvo EO, Ferreira RR, Madeira FS, Alves GF, Chambela MC, Mendes VG, et al. Correlation of transforming growth factor- β 1 and tumour necrosis factor levels with left ventricular function in Chagas disease. *Mem Inst Oswaldo Cruz.* 19 févr 2018;113(4):e170440.
248. Antúnez MI, Cardoni RL. IL-12 and IFN-gamma production, and NK cell activity, in acute and chronic experimental *Trypanosoma cruzi* infections. *Immunol Lett.* 1 févr 2000;71(2):103-9.
249. Soares MBP, de Lima RS, Rocha LL, Vasconcelos JF, Rogatto SR, dos Santos RR, et al. Gene expression changes associated with myocarditis and fibrosis in hearts of mice with chronic chagasic cardiomyopathy. *J Infect Dis.* 15 août 2010;202(3):416-26.
250. Rodrigues AA, Notário AFO, Teixeira TL, Silva RT e, Quintal APN, Alves RN, et al. A high throughput analysis of cytokines and chemokines expression during the course of *Trypanosoma cruzi* experimental oral infection. *Acta Tropica.* 1 mai 2016;157:42-53.
251. do Carmo Neto JR, Vinicius da Silva M, Braga YLL, Florencio da Costa AW, Fonseca SG, Nagib PRA, et al. Correlation between intestinal BMP2, IFN γ , and neural death in experimental infection with *Trypanosoma cruzi*. *PLoS One.* 9 févr 2021;16(2):e0246692.

252. Wen JJ, Gupta S, Guan Z, Dhiman M, Condon D, Lui C, et al. Phenyl- α -tert-butyl-nitrone and Benzonidazole Treatment Controlled the Mitochondrial Oxidative Stress and Evolution of Cardiomyopathy in Chronic Chagasic Rats. *J Am Coll Cardiol*. 1 juin 2010;55(22):2499-508.
253. Guedes PMM, Veloso VM, Afonso LCC, Caliari MV, Carneiro CM, Diniz LF, et al. Development of chronic cardiomyopathy in canine Chagas disease correlates with high IFN- γ , TNF- α , and low IL-10 production during the acute infection phase. *Veterinary Immunology and Immunopathology*. 15 juill 2009;130(1):43-52.
254. Wu J, Cao J, Fan Y, Li C, Hu X. Comprehensive analysis of miRNA–mRNA regulatory network and potential drugs in chronic chagasic cardiomyopathy across human and mouse. *BMC Med Genomics*. 29 nov 2021;14:283.
255. Zhou L, Li Z, Li J, Yang S, Gong H. Detecting imperative genes and infiltrating immune cells in chronic Chagas cardiomyopathy by bioinformatics analysis. *Infection, Genetics and Evolution*. 1 nov 2021;95:105079.
256. Gómez I, López MC, Egui A, Palacios G, Carrilero B, Benítez C, et al. Differential expression profile of genes involved in the immune response associated to progression of chronic Chagas disease. *PLoS Negl Trop Dis*. 13 juill 2023;17(7):e0011474.
257. El-Hattab AW, Scaglia F. Mitochondrial Cardiomyopathies. *Front Cardiovasc Med*. 25 juill 2016;3:25.
258. Oliveira TB de, Pedrosa RC, Filho DW. Oxidative stress in chronic cardiopathy associated with Chagas disease. *International Journal of Cardiology*. 4 avr 2007;116(3):357-63.
259. Teixeira PC, Ducret A, Langen H, Nogoceke E, Santos RHB, Silva Nunes JP, et al. Impairment of Multiple Mitochondrial Energy Metabolism Pathways in the Heart of Chagas Disease Cardiomyopathy Patients. *Front Immunol*. 12 nov 2021;12:755782.
260. Laugier L, Frade AF, Ferreira FM, Baron MA, Teixeira PC, Cabantous S, et al. Whole-Genome Cardiac DNA Methylation Fingerprint and Gene Expression Analysis Provide New Insights in the Pathogenesis of Chronic Chagas Disease Cardiomyopathy. *Clin Infect Dis*. 1 oct 2017;65(7):1103-11.
261. Wen JJ, Zago MP, Nuñez S, Gupta S, Burgos FN, Garg NJ. Serum Proteomic Signature of Human Chagasic Patients for the Identification of Novel Potential Protein Biomarkers of Disease. *Mol Cell Proteomics*. août 2012;11(8):435-52.
262. do Carmo Neto JR, da Costa AWF, Braga YLL, Lucio FH, dos Santos Martins ALM, dos Reis MA, et al. The Colombian Strain of *Trypanosoma cruzi* Induces a Proinflammatory Profile, Neuronal Death, and Collagen Deposition in the Intestine of C57BL/6 Mice Both during the Acute and Early Chronic Phase. *Mediators Inflamm*. 12 janv 2022;2022:7641357.
263. Barbosa-Ferreira JM, Mady C, Ianni BM, Lopes HF, Ramires FJA, Salemi VMC, et al. Dysregulation of Autonomic Nervous System in Chagas' Heart Disease Is Associated with Altered Adipocytokines Levels. *PLoS One*. 6 juill 2015;10(7):e0131447.
264. Casares-Marfil D, Kerick M, Andrés-León E, Bosch-Nicolau P, Molina I, Martin J, et al. GWAS loci associated with Chagas cardiomyopathy influences DNA methylation levels. *PLoS Negl Trop Dis*. 29 oct 2021;15(10):e0009874.
265. Nonaka CKV, Macêdo CT, Cavalcante BRR, de Alcântara AC, Silva DN, Bezerra M da R, et al. Circulating miRNAs as Potential Biomarkers Associated with Cardiac Remodeling and Fibrosis in Chagas Disease Cardiomyopathy. *Int J Mol Sci*. 20 août 2019;20(16):4064.
266. Navarro IC, Ferreira FM, Nakaya HI, Baron MA, Vilar-Pereira G, Pereira IR, et al. MicroRNA Transcriptome Profiling in Heart of *Trypanosoma cruzi*-Infected Mice: Parasitological and Cardiological Outcomes. *PLoS Negl Trop Dis*. 18 juin 2015;9(6):e0003828.

267. Ferreira LRP, Ferreira FM, Laugier L, Cabantous S, Navarro IC, da Silva Cândido D, et al. Integration of miRNA and gene expression profiles suggest a role for miRNAs in the pathobiological processes of acute *Trypanosoma cruzi* infection. *Sci Rep*. 21 déc 2017;7:17990.
268. Jha BK, Varikuti S, Seidler GR, Volpedo G, Satoskar AR, McGwire BS. MicroRNA-155 Deficiency Exacerbates *Trypanosoma cruzi* Infection. *Infect Immun*. 22 juin 2020;88(7):e00948-19.
269. Frade AF, Laugier L, Ferreira LRP, Baron MA, Benvenuti LA, Teixeira PC, et al. Myocardial Infarction–Associated Transcript, a Long Noncoding RNA, Is Overexpressed During Dilated Cardiomyopathy Due to Chronic Chagas Disease. *The Journal of Infectious Diseases*. 1 juill 2016;214(1):161-5.
270. Krassowski M, Das V, Sahu SK, Misra BB. State of the Field in Multi-Omics Research: From Computational Needs to Data Mining and Sharing. *Front Genet*. 10 déc 2020;11:610798.
271. Aksenov SV, Church B, Dhiman A, Georgieva A, Sarangapani R, Helmlinger G, et al. An integrated approach for inference and mechanistic modeling for advancing drug development. *FEBS Letters*. 2005;579(8):1878-83.
272. Minagawa H, Yamashita T, Honda M, Tabuse Y, Kamijo K, Tsugita A, et al. Comparative Analysis of Proteome and Transcriptome in Human Hepatocellular Carcinoma using 2D-DIGE and SAGE. *Protein J*. 1 déc 2008;27(7):409-19.
273. Kotliarov Y, Kotliarova S, Charong N, Li A, Walling J, Aquilanti E, et al. Correlation Analysis between SNP and Expression Arrays in Gliomas Identify Potentially Relevant Targets Genes. *Cancer Res*. 15 févr 2009;69(4):1596-603.
274. de Anda-Jáuregui G, Hernández-Lemus E. Computational Oncology in the Multi-Omics Era: State of the Art. *Front Oncol*. 7 avr 2020;10:423.
275. Enjalbert B, Jourdan F, Portais JC. Intuitive Visualization and Analysis of Multi-Omics Data and Application to *Escherichia coli* Carbon Metabolism. *PLoS One*. 22 juin 2011;6(6):e21318.
276. Droste P, Miebach S, Niedenführ S, Wiechert W, Nöh K. Visualizing multi-omics data in metabolic networks with the software Omix—A case study. *Biosystems*. 1 août 2011;105(2):154-61.
277. Gibbs DL, Gralinski L, Baric RS, McWeeney SK. Multi-omic network signatures of disease. *Front Genet*. 7 janv 2014;4:309.
278. Bodein A, Scott-Boyer MP, Perin O, Lê Cao KA, Droit A. Interpretation of network-based integration from multi-omics longitudinal data. *Nucleic Acids Res*. 9 déc 2021;50(5):e27.
279. Argelaguet R, Velten B, Arnol D, Dietrich S, Zenz T, Marioni JC, et al. Multi-Omics Factor Analysis—a framework for unsupervised integration of multi-omics data sets. *Mol Syst Biol*. 20 juin 2018;14(6):e8124.
280. Picard M, Scott-Boyer MP, Bodein A, Périn O, Droit A. Integration strategies of multi-omics data for machine learning analysis. *Comput Struct Biotechnol J*. 22 juin 2021;19:3735-46.
281. Vaske CJ, Benz SC, Sanborn JZ, Earl D, Szeto C, Zhu J, et al. Inference of patient-specific pathway activities from multi-dimensional cancer genomics data using PARADIGM. *Bioinformatics*. 15 juin 2010;26(12):i237-45.
282. Schaefer CF, Anthony K, Krupa S, Buchoff J, Day M, Hannay T, et al. PID: the Pathway Interaction Database. *Nucleic Acids Res*. janv 2009;37(Database issue):D674-9.
283. Si H, Lu H, Yang X, Mattox A, Jang M, Bian Y, et al. TNF- α modulates genome-wide redistribution of Δ Np63 α /TAp73 and NF- κ B c-REL interactive binding on TP53 and AP-1 motifs to promote an oncogenic gene program in squamous cancer. *Oncogene*. 3 nov 2016;35(44):5781-94.
284. Koh HWL, Fermin D, Vogel C, Choi KP, Ewing RM, Choi H. iOmicsPASS: network-based integration of multiomics data for predictive subnetwork discovery. *NPJ Syst Biol Appl*. 9 juill 2019;5:22.

285. Koh HWL, Pilbrow AP, Tan SH, Zhao Q, Benke PI, Burla B, et al. An integrated signature of extracellular matrix proteins and a diastolic function imaging parameter predicts post-MI long-term outcomes. *Front Cardiovasc Med*. 12 avr 2023;10:1123682.
286. Lock EF, Hoadley KA, Marron JS, Nobel AB. JOINT AND INDIVIDUAL VARIATION EXPLAINED (JIVE) FOR INTEGRATED ANALYSIS OF MULTIPLE DATA TYPES. *Ann Appl Stat*. 1 mars 2013;7(1):523-42.
287. Meng C, Kuster B, Culhane AC, Gholami AM. A multivariate approach to the integration of multi-omics datasets. *BMC Bioinformatics*. 29 mai 2014;15:162.
288. Argelaguet R, Arnol D, Bredikhin D, Deloro Y, Velten B, Marioni JC, et al. MOFA+: a statistical framework for comprehensive integration of multi-modal single-cell data. *Genome Biol*. 11 mai 2020;21:1111.
289. Singh A, Shannon CP, Gautier B, Rohart F, Vacher M, Tebbutt SJ, et al. DIABLO: an integrative approach for identifying key molecular drivers from multi-omics assays. *Bioinformatics*. 1 sept 2019;35(17):3055-62.
290. Haas EA, Saad MJA, Santos A, Vitulo N, Lemos WJF, Martins AMA, et al. A red wine intervention does not modify plasma trimethylamine N-oxide but is associated with broad shifts in the plasma metabolome and gut microbiota composition. *Am J Clin Nutr*. 7 oct 2022;116(6):1515-29.
291. Ganugi P, Fiorini A, Ardenti F, Caffi T, Bonini P, Taskin E, et al. Nitrogen use efficiency, rhizosphere bacterial community, and root metabolome reprogramming due to maize seed treatment with microbial biostimulants. *Physiol Plant*. 2022;174(2):e13679.
292. Francavilla A, Ferrero G, Pardini B, Tarallo S, Zanatto L, Caviglia GP, et al. Gluten-free diet affects fecal small non-coding RNA profiles and microbiome composition in celiac disease supporting a host-gut microbiota crosstalk. *Gut Microbes*. 15(1):2172955.
293. Zhao Y, Klein A, Castellanos FX, Milham MP. Brain Age Prediction: Cortical and Subcortical Shape Covariation in the Developing Human Brain. *Neuroimage*. 15 nov 2019;202:116149.
294. Zhao Y, Paulus M, Bagot KS, Constable RT, Yaggi HK, Redeker NS, et al. Brain structural covariation linked to screen media activity and externalizing behaviors in children. *J Behav Addict*. 30 juin 2022;11(2):417-26.
295. Leng D, Zheng L, Wen Y, Zhang Y, Wu L, Wang J, et al. A benchmark study of deep learning-based multi-omics data fusion methods for cancer. *Genome Biol*. 9 août 2022;23:171.
296. Wekesa JS, Kimwele M. A review of multi-omics data integration through deep learning approaches for disease diagnosis, prognosis, and treatment. *Front Genet*. 20 juill 2023;14:1199087.
297. Dong X, Lin L, Zhang R, Zhao Y, Christiani DC, Wei Y, et al. TOBML: trans-omics block missing data imputation using a k-nearest neighbor weighted approach. *Bioinformatics*. 15 avr 2019;35(8):1278-83.
298. Voillet V, Besse P, Liaubet L, San Cristobal M, González I. Handling missing rows in multi-omics data integration: multiple imputation in multiple factor analysis framework. *BMC Bioinformatics*. 3 oct 2016;17:402.
299. Escofier B, Pagès J. Multiple factor analysis (AFMULT package). *Computational Statistics & Data Analysis*. 1 août 1994;18(1):121-40.
300. Chen L, Xu J, Li SC. DeepMF: deciphering the latent patterns in omics profiles with a deep learning method. *BMC Bioinformatics*. 27 déc 2019;20(Suppl 23):648.
301. Tarazona S, Balzano-Nogueira L, Gómez-Cabrero D, Schmidt A, Imhof A, Hankemeier T, et al. Harmonization of quality metrics and power calculation in multi-omic studies. *Nat Commun*. 18 juin 2020;11:3092.
302. Zhao Z, Anand R, Wang M. Maximum Relevance and Minimum Redundancy Feature Selection Methods for a Marketing Machine Learning Platform. In: 2019 IEEE International

- Conference on Data Science and Advanced Analytics (DSAA) [Internet]. 2019 [cité 5 oct 2023]. p. 442-52. Disponible sur: <https://ieeexplore.ieee.org/document/8964172>
303. Zeng X, Chen YW, Tao C. Feature Selection Using Recursive Feature Elimination for Handwritten Digit Recognition. In: 2009 Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing [Internet]. 2009 [cité 5 oct 2023]. p. 1205-8. Disponible sur: <https://ieeexplore.ieee.org/document/5337549>
304. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. *J Artif Int Res.* 1 juin 2002;16(1):321-57.
305. He H, Bai Y, Garcia EA, Li S. ADASYN: Adaptive synthetic sampling approach for imbalanced learning. In: 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence) [Internet]. 2008 [cité 5 oct 2023]. p. 1322-8. Disponible sur: <https://ieeexplore.ieee.org/document/4633969>
306. Hira MT, Razzaque MA, Angione C, Scrivens J, Sawan S, Sarker M. Integrated multi-omics analysis of ovarian cancer using variational autoencoders. *Sci Rep.* 18 mars 2021;11:6265.
307. Islam A, Belhaouari SB, Rehman AU, Bensmail H. KNNOR: An oversampling technique for imbalanced datasets. *Applied Soft Computing.* 1 janv 2022;115:108288.
308. Figueira A, Vaz B. Survey on Synthetic Data Generation, Evaluation Methods and GANs. *Mathematics.* janv 2022;10(15):2733.
309. Huang S, Chaudhary K, Garmire LX. More Is Better: Recent Progress in Multi-Omics Data Integration Methods. *Front Genet.* 16 juin 2017;8:84.
310. Subramanian I, Verma S, Kumar S, Jere A, Anamika K. Multi-omics Data Integration, Interpretation, and Its Application. *Bioinform Biol Insights.* 31 janv 2020;14:1177932219899051.
311. Athieniti E, Spyrou GM. A guide to multi-omics data collection and integration for translational medicine. *Comput Struct Biotechnol J.* 1 déc 2022;21:134-49.
312. Love M. awesome-multi-omics [Internet]. 2023 [cité 5 oct 2023]. Disponible sur: <https://github.com/mikelove/awesome-multi-omics>
313. Cantini L, Zakeri P, Hernandez C, Naldi A, Thieffry D, Remy E, et al. Benchmarking joint multi-omics dimensionality reduction approaches for the study of cancer. *Nat Commun.* 5 janv 2021;12:124.
314. Strefeler A, Jan M, Quadroni M, Teav T, Rosenberg N, Chatton JY, et al. Molecular insights into sex-specific metabolic alterations in Alzheimer's mouse brain using multi-omics approach. *Alzheimers Res Ther.* 9 janv 2023;15:8.
315. Sudhakar P, Verstockt B, Cremer J, Verstockt S, Sabino J, Ferrante M, et al. Understanding the Molecular Drivers of Disease Heterogeneity in Crohn's Disease Using Multi-omic Data Integration and Network Analysis. *Inflamm Bowel Dis.* 14 déc 2020;27(6):870-86.
316. Bhatt RR, Gupta A, Labus JS, Liu C, Vora PP, Jean Stains, et al. A neuropsychosocial signature predicts longitudinal symptom changes in women with irritable bowel syndrome. *Mol Psychiatry.* 2022;27(3):1774-91.
317. Laugier L, Ferreira LRP, Ferreira FM, Cabantous S, Frade AF, Nunes JP, et al. miRNAs may play a major role in the control of gene expression in key pathobiological processes in Chagas disease cardiomyopathy. *PLoS Negl Trop Dis.* 22 déc 2020;14(12):e0008889.
318. Brochet P, Ianni BM, Laugier L, Frade AF, Silva Nunes JP, Teixeira PC, et al. Epigenetic regulation of transcription factor binding motifs promotes Th1 response in Chagas disease cardiomyopathy. *Front Immunol.* 22 août 2022;13:958200.
319. Hogan PG. Calcium–NFAT transcriptional signalling in T cell activation and T cell exhaustion. *Cell Calcium.* mai 2017;63:66-9.
320. Mutlu U, Cakmak R, Sonsöz MR, Karaayvaz EB, Uzum AK, Tanakol R, et al. A rare cause of dilated cardiomyopathy: hypocalcemia. *Arch Endocrinol Metab.* 8 sept 2022;66(4):582-6.

321. Teixeira PC, Iwai LK, Kuramoto ACK, Honorato R, Fiorelli A, Stolf N, et al. Proteomic inventory of myocardial proteins from patients with chronic Chagas' cardiomyopathy. *Braz J Med Biol Res.* déc 2006;39(12):1549-62.
322. Corsten MF, Papageorgiou A, Verhesen W, Carai P, Lindow M, Obad S, et al. MicroRNA profiling identifies microRNA-155 as an adverse mediator of cardiac injury and dysfunction during acute viral myocarditis. *Circ Res.* 3 août 2012;111(4):415-25.
323. Brochet P, Ianni B, Nunes JPS, Frade AF, Teixeira PC, Mady C, et al. Blood DNA methylation marks discriminate Chagas cardiomyopathy disease clinical forms. *Front Immunol.* 29 sept 2022;13:1020572.
324. Dhar GA, Saha S, Mitra P, Nag Chaudhuri R. DNA methylation and regulation of gene expression: Guardian of our health. *Nucleus (Calcutta).* 2021;64(3):259-70.
325. Fernández-Caggiano M, Barallobre-Barreiro J, Rego-Pérez I, Crespo-Leiro MG, Paniagua MJ, Grillé Z, et al. Mitochondrial DNA haplogroup H as a risk factor for idiopathic dilated cardiomyopathy in Spanish population. *Mitochondrion.* juill 2013;13(4):263-8.
326. Larsen S, Díez-Sánchez C, Rabøl R, Ara I, Dela F, Helge JW. Increased intrinsic mitochondrial function in humans with mitochondrial haplogroup H. *Biochim Biophys Acta.* févr 2014;1837(2):226-31.
327. Nunes JPS, Andrieux P, Brochet P, Almeida RR, Kitano E, Honda AK, et al. Co-Exposure of Cardiomyocytes to IFN- γ and TNF- α Induces Mitochondrial Dysfunction and Nitro-Oxidative Stress: Implications for the Pathogenesis of Chronic Chagas Disease Cardiomyopathy. *Front Immunol.* 11 nov 2021;12:755862.
328. Walker BR, Moraes CT. Nuclear-Mitochondrial Interactions. *Biomolecules.* 10 mars 2022;12(3):427.
329. Kingdom R, Wright CF. Incomplete Penetrance and Variable Expressivity: From Clinical Studies to Population Cohorts. *Front Genet.* 25 juill 2022;13:920390.
330. Silva KDA, Nunes JPS, Andrieux P, Brochet P, Almeida RR, Kuramoto Takara ACK, et al. Chagas Disease Megaesophagus Patients Carrying Variant MRPS18B P260A Display Nitro-Oxidative Stress and Mitochondrial Dysfunction in Response to IFN- γ Stimulus. *Biomedicines.* 7 sept 2022;10(9):2215.
331. Brochet P, Mouren JC, Hannouche L, Lopez F, Ballester B, Cunha-Neto E, et al. ChagasDB: 80 years of publicly available data on the molecular host response to *Trypanosoma cruzi* infection in a single database. *Database (Oxford).* 26 mai 2023;2023:baad037.
332. Stein CM, Weiskirchen R, Damm F, Strzelecka PM. Single-cell omics: Overview, analysis, and application in biomedical science. *J Cell Biochem.* nov 2021;122(11):1571-8.
333. Koenig AL, Shchukina I, Amrute J, Andhey PS, Zaitsev K, Lai L, et al. Single-cell transcriptomics reveals cell-type-specific diversification in human heart failure. *Nat Cardiovasc Res.* mars 2022;1(3):263-80.
334. Lareau CA, Ludwig LS, Muus C, Gohil SH, Zhao T, Chiang Z, et al. Massively parallel single-cell mitochondrial DNA genotyping and chromatin profiling. *Nat Biotechnol.* avr 2021;39(4):451-61.
335. Ferragut F, Acevedo GR, Gómez KA. T Cell Specificity: A Great Challenge in Chagas Disease. *Front Immunol.* 29 juin 2021;12:674078.
336. Regev A, Teichmann SA, Lander ES, Amit I, Benoist C, Birney E, et al. The Human Cell Atlas. *eLife.* 6:e27041.
337. Djuretic IM, Levanon D, Negreanu V, Groner Y, Rao A, Ansel KM. Transcription factors T-bet and Runx3 cooperate to activate *Irf4* and silence *Irf4* in T helper type 1 cells. *Nat Immunol.* févr 2007;8(2):145-53.

338. Blagus R, Lusa L. SMOTE for high-dimensional class-imbalanced data. *BMC Bioinformatics*. 22 mars 2013;14:106.
339. Shukla V, Lu R. IRF4 and IRF8: Governing the virtues of B Lymphocytes. *Front Biol (Beijing)*. août 2014;9(4):269-82.
340. Huber M, Lohoff M. IRF4 at the crossroads of effector T-cell fate decision. *Eur J Immunol*. juill 2014;44(7):1886-95.
341. Zheng Y, Chaudhry A, Kas A, deRoos P, Kim JM, Chu TT, et al. Regulatory T-cell suppressor program co-opts transcription factor IRF4 to control TH2 responses. *Nature*. 19 mars 2009;458(7236):351-6.
342. Qiu Y, Xu J, Yang L, Zhao G, Ding J, Chen Q, et al. MiR-375 silencing attenuates pro-inflammatory macrophage response and foam cell formation by targeting KLF4. *Exp Cell Res*. 1 mars 2021;400(1):112507.
343. Kinjyo I, Inoue H, Hamano S, Fukuyama S, Yoshimura T, Koga K, et al. Loss of SOCS3 in T helper cells resulted in reduced immune responses and hyperproduction of interleukin 10 and transforming growth factor- β 1. *J Exp Med*. 17 avr 2006;203(4):1021-31.
344. Alam Z, Devalaraja S, Li M, To TKJ, Folkert IW, Mitchell-Velasquez E, et al. Counter Regulation of Spic by NF- κ B and STAT Signaling Controls Inflammation and Iron Metabolism in Macrophages. *Cell Rep*. 30 juin 2020;31(13):107825.
345. Davila A, Xu Z, Li S, Rozewicki J, Wilamowski J, Kotelnikov S, et al. AbAdapt: an adaptive approach to predicting antibody-antigen complex structures from sequence. *Bioinform Adv*. 7 mars 2022;2(1):vbac015.
346. He J, Shen J, Luo W, Han Z, Xie F, Pang T, et al. Research progress on application of single-cell TCR/BCR sequencing technology to the tumor immune microenvironment, autoimmune diseases, and infectious diseases. *Front Immunol*. 18 août 2022;13:969808.
347. Wang J, Chen X, Huang W. MicroRNA-369 attenuates hypoxia-induced cardiomyocyte apoptosis and inflammation via targeting TRPV3. *Braz J Med Biol Res*. 15 janv 2021;54(3):e10550.
348. Scalavino V, Liso M, Cavalcanti E, Gigante I, Lippolis A, Mastronardi M, et al. miR-369-3p modulates inducible nitric oxide synthase and is involved in regulation of chronic inflammatory response. *Sci Rep*. 29 sept 2020;10(1):15942.
349. Matchett KB, Lynam-Lennon N, Watson RW, Brown JAL. Advances in Precision Medicine: Tailoring Individualized Therapies. *Cancers (Basel)*. 25 oct 2017;9(11):146.

Annexes

J'ai choisi de placer en annexe deux articles auxquels j'ai contribué, car ils ne correspondent pas directement à mon projet de thèse, mais apportent néanmoins des résultats intéressants et des informations complémentaires sur la maladie de Chagas.

Annexe 1

Co-Exposure of Cardiomyocytes to IFN- γ and TNF- α Induces Mitochondrial Dysfunction and Nitro-Oxidative Stress: Implications for the Pathogenesis of Chronic Chagas Disease Cardiomyopathy

João Paulo Silva Nunes*, Pauline Andrieux*, **Pauline Brochet**, Rafael Ribeiro Almeida, Eduardo Kitano, André Kenji Honda, Leo Kei Iwai, Débora Andrade-Silva, David Goudenège, Karla Deysiree Alcântara Silva, Raquel de Souza Vieira, Débora Levy, Sergio Paulo Bydlowski, Frédéric Gallardo, Magali Torres, Edimar Alcides Bocchi, Miguel Mano, Ronaldo Honorato Barros Santos, Fernando Bacal, Pablo Pomerantzeff, Francisco Rafael Martins Laurindo, Priscila Camillo Teixeira, Helder I. Nakaya, Jorge Kalil, Vincent Procaccio, Christophe Chevillard*, Edecio Cunha-Neto*.

Front Immunol. 2021 Nov 11;12:755862. doi: 10.3389/fimmu.2021.755862. PMID: 34867992; PMCID: PMC8632642.

Présentation de la publication :

Le développement des CCC a été associé à une forte production de TNF- α et d'IFN- γ , ainsi qu'à des dysfonctions mitochondriales. Lors de l'infection par *T. cruzi*, ces deux cytokines sont connues pour induire la production de ROS par les mitochondries, conduisant à la destruction du parasite. Lors de cette étude, nous avons voulu évaluer l'impact du TNF- α et de l'IFN- γ sur les mitochondries cardiaques, en stimulant des cellules AC-16 in vitro avec ces cytokines.

Pour confirmer l'importance des dysfonctionnements mitochondriaux dans la pathogénèse des CCC, la quantité de mitochondries présentes dans le cœur de patients atteints de CCC a été quantifiée à l'aide de gènes marqueurs, de même que la concentration en NO. Ces valeurs ont alors été comparées à celles mesurées chez des patients atteints de CMD. Il a été constaté que le nombre de mitochondries cardiaques était significativement plus faible chez les patients atteints de CCC que chez les CMD et, à l'inverse, la concentration en NO était significativement plus élevée.

La stimulation des cellules AC-16 par de l'IFN- γ a induit une diminution significative du potentiel de membrane mitochondrial et de la production d'ATP, ainsi qu'une augmentation de la production de NO. Ce phénomène a été plus important lors de la stimulation des cellules à la fois par de l'IFN- γ et du TNF- α . De plus, 275 protéines ont été identifiées comme différentiellement exprimés par les AC-16 avant et après stimulation, dont 37 impliquées dans des processus mitochondriaux. Parmi les 19 protéines qui ont été quantifiées à la fois dans cette analyse et dans une étude protéomique réalisée sur des tissus cardiaques (259), 9 présentent des variations d'expression communes, ce qui confirme des similitudes entre le développement des CCC et la réponse des mitochondries à ces deux cytokines.

Afin de simuler l'infection par *T. cruzi* et le développement des CCC, des cellules AC-16 ont été stimulées avec de l'IFN- γ et du TNF- α pendant 1 à 48 heures, et un séquençage de l'ARN a été réalisé à différents temps. Au total, 1.443 gènes ont été identifiés comme différentiellement exprimés, majoritairement associés à la réponse immunitaire, à l'inflammation et aux maladies cardiaques. De plus, parmi tous les gènes différentiellement exprimés identifiés, 67 sont associés aux processus mitochondriaux.

Pour finir, les cellules AC-16, après stimulation par du TNF- α et/ou de l'IFN- γ , ont été traitées par différentes molécules régulant (positivement ou négativement) des protéines impliquées dans la production de ROS/RNS. 12 d'entre elles ont conduit à une augmentation significative du potentiel de membranes des AC-16.

Pour conclure, cette étude a permis de confirmer la pertinence de la stimulation des cellules AC-16 par le TNF- α et/ou l'IFN- γ en tant que modèle d'étude des CCC. Elle a également mis en évidence des gènes et des protéines affectés par cette stimulation. De plus, nous avons identifié 12 molécules restaurant significativement la fonction mitochondriale. Des études complémentaires seront nécessaires afin de confirmer leur efficacité in vivo, et leur éventuelle application dans le traitement des CCC.

Contribution personnelle :

J'ai effectué l'ensemble des analyses bio-informatiques de cette étude, comprenant le traitement des données de séquençage et les analyses statistiques en résultant. J'ai également participé à la préparation et aux corrections de cet article, dont je suis 3^{ème} auteur.



OPEN ACCESS

Edited by:

Clarissa M. Maya-Monteiro,
Oswaldo Cruz Foundation (Fiocruz),
Brazil

Reviewed by:

Emiliano Medei,
Federal University of Rio de Janeiro,
Brazil

Joseli Lannes-Vieira,
Oswaldo Cruz Foundation (Fiocruz),
Brazil

Claudia Paiva,
Federal University of Rio de Janeiro,
Brazil

*Correspondence:

Edecio Cunha-Neto
edecunha@gmail.com
Christophe Chevillard
christophe.chevillard@univ-amu.fr

†These authors have contributed
equally to this work

‡These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Inflammation,
a section of the journal
Frontiers in Immunology

Received: 09 August 2021

Accepted: 11 October 2021

Published: 11 November 2021

Citation:

Nunes JPS, Andrieux P, Brochet P,
Almeida RR, Kitano E, Honda AK,
Iwai LK, Andrade-Silva D, Goudenège D,
Alcântara Silva KD, Vieira RdS, Levy D,
Bydlowski SP, Gallardo F, Torres M,
Bocchi EA, Mano M, Santos RHB,
Bacal F, Pomerantzeff P, Laurindo FRM,
Teixeira PC, Nakaya HI, Kalil J,
Procaccio V, Chevillard C and
Cunha-Neto E (2021) Co-Exposure of
Cardiomyocytes to IFN- γ and TNF- α
Induces Mitochondrial Dysfunction and
Nitro-Oxidative Stress: Implications for
the Pathogenesis of Chronic Chagas
Disease Cardiomyopathy.
Front. Immunol. 12:755862.
doi: 10.3389/fimmu.2021.755862

Co-Exposure of Cardiomyocytes to IFN- γ and TNF- α Induces Mitochondrial Dysfunction and Nitro-Oxidative Stress: Implications for the Pathogenesis of Chronic Chagas Disease Cardiomyopathy

João Paulo Silva Nunes^{1,2,3,4†}, Pauline Andrieux^{4†}, Pauline Brochet⁴, Rafael Ribeiro Almeida^{1,3}, Eduardo Kitano¹, André Kenji Honda¹, Leo Kei Iwai⁵, Débora Andrade-Silva⁵, David Goudenège⁶, Karla Deysiree Alcântara Silva^{1,2}, Raquel de Souza Vieira¹, Débora Levy¹, Sergio Paulo Bydlowski¹, Frédéric Gallardo⁴, Magali Torres⁴, Edimar Alcides Bocchi⁷, Miguel Mano⁸, Ronaldo Honorato Barros Santos⁹, Fernando Bacal⁹, Pablo Pomerantzeff⁹, Francisco Rafael Martins Laurindo¹⁰, Priscila Camillo Teixeira¹¹, Helder I. Nakaya¹¹, Jorge Kalil^{1,2,3}, Vincent Procaccio¹³, Christophe Chevillard^{4*†} and Edecio Cunha-Neto^{1,2,3*†}

¹ Laboratory of Immunology, Heart Institute (Incor), Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, São Paulo, Brazil, ² Division of Clinical Immunology and Allergy, Faculdade de Medicina da Universidade de São Paulo, São Paulo, Brazil, ³ iii-Institute for Investigation in Immunology, Instituto Nacional de Ciência e Tecnologia (INCT), São Paulo, Brazil, ⁴ INSERM, UMR_1090, Aix Marseille Université, TAGC Theories and Approaches of Genomic Complexity, Institut MarMaRa, Marseille, France, ⁵ Laboratório Especial de Toxinologia Aplicada, Center of Toxins, Immune-Response and Cell Signaling (CeTICS), Instituto Butantan, São Paulo, Brazil, ⁶ Department of Biochemistry and Genetics, University Hospital of Angers, Angers, France, ⁷ Heart Failure Team, Heart Institute (Incor) Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, São Paulo, Brazil, ⁸ Functional Genomics and RNA-based Therapeutics Laboratory, Center for Neuroscience and Cell Biology (CNC), University of Coimbra, Coimbra, Portugal, ⁹ Division of Surgery, Heart Institute, School of Medicine, University of São Paulo, São Paulo, Brazil, ¹⁰ Laboratory of Vascular Biology, Heart Institute of the School of Medicine, University of São Paulo, São Paulo, Brazil, ¹¹ Translational Research Sciences, Pharma Research and Early Development F. Hoffmann-La Roche, Basel, Switzerland, ¹² Hospital Israelita Albert Einstein, São Paulo, Brazil, ¹³ MitoLab, UMR CNRS 6015-INSERM U1083, Université d'Angers, Angers, France

Infection by the protozoan *Trypanosoma cruzi* causes Chagas disease cardiomyopathy (CCC) and can lead to arrhythmia, heart failure and death. Chagas disease affects 8 million people worldwide, and chronic production of the cytokines IFN- γ and TNF- α by T cells together with mitochondrial dysfunction are important players for the poor prognosis of the disease. Mitochondria occupy 40% of the cardiomyocytes volume and produce 95% of cellular ATP that sustain the life-long cycles of heart contraction. As IFN- γ and TNF- α have been described to affect mitochondrial function, we hypothesized that IFN- γ and TNF- α are involved in the myocardial mitochondrial dysfunction observed in CCC patients. In this study, we quantified markers of mitochondrial dysfunction and nitro-oxidative stress in CCC heart tissue and in IFN- γ /TNF- α -stimulated AC-16 human cardiomyocytes. We found that CCC myocardium displayed increased levels of nitro-oxidative stress and reduced mitochondrial DNA as compared with myocardial tissue from patients with dilated cardiomyopathy (DCM). IFN- γ /TNF- α treatment of AC-16 cardiomyocytes induced increased nitro-oxidative stress and decreased the mitochondrial membrane potential

($\Delta\Psi_m$). We found that the STAT1/NF- κ B/NOS2 axis is involved in the IFN- γ /TNF- α -induced decrease of $\Delta\Psi_m$ in AC-16 cardiomyocytes. Furthermore, treatment with mitochondria-sparing agonists of AMPK, NRF2 and SIRT1 rescues $\Delta\Psi_m$ in IFN- γ /TNF- α -stimulated cells. Proteomic and gene expression analyses revealed that IFN- γ /TNF- α -treated cells corroborate mitochondrial dysfunction, transmembrane potential of mitochondria, altered fatty acid metabolism and cardiac necrosis/cell death. Functional assays conducted on Seahorse respirometer showed that cytokine-stimulated cells display decreased glycolytic and mitochondrial ATP production, dependency of fatty acid oxidation as well as increased proton leak and non-mitochondrial oxygen consumption. Together, our results suggest that IFN- γ and TNF- α cause direct damage to cardiomyocytes' mitochondria by promoting oxidative and nitrosative stress and impairing energy production pathways. We hypothesize that treatment with agonists of AMPK, NRF2 and SIRT1 might be an approach to ameliorate the progression of Chagas disease cardiomyopathy.

Keywords: mitochondrial dysfunction, chronic Chagas disease cardiomyopathy, interferon gamma, energy metabolism, mitochondria

INTRODUCTION

Heart failure (HF) is an important worldwide public health problem. Available therapies are insufficient and do not fully address molecular abnormalities that occur in cardiomyocytes (1). Chagas disease cardiomyopathy (CCC) accounts for 25% of HF cases and is a major cause of death in Latin America (2, 3). CCC is a severe inflammatory dilated cardiomyopathy caused by persistent infection by the protozoan *Trypanosoma cruzi*. While 60% of Chagas Disease (CD) patients are mostly asymptomatic in the so-called “indeterminate” form (IF) and do not develop heart disease, CCC patients (30%, roughly 8 million people) display HF, arrhythmia, and disability (4, 5). CCC patients have 50% shorter survival rate and worse prognosis compared to patients with cardiomyopathies of non-inflammatory etiologies, such as ischemic, idiopathic and hypertensive cardiomyopathies (2).

The pathogenesis of CCC is still to be completely understood (6). Low-grade chronic *T. cruzi* persistence, which drives continued production of IFN- γ and TNF- α by Th1 T cells immune cells, plays a central pathogenic role in CCC (2, 4, 5). Indeed, omics and immunohistochemistry studies revealed higher levels of IFN- γ and TNF- α in heart tissue from CCC patients compared to other cardiomyopathies (7–11). Activation of the NF- κ B/NOS2 axis by the continued production of IFN- γ /TNF- α can promote increased levels of intracellular reactive nitrogen species (RNS) which are important for parasite control (12, 13). However, long-term sustained activation of this axis may promote damaging accumulation of reactive oxygen species (ROS) through induction of NADP oxidases and overproduction of mitochondrial reactive oxygen species (ROS), such as hydrogen peroxide (H₂O₂) and superoxide anion (O₂⁻) and RNS, molecules widely known to induce mitochondrial dysfunction and disturbances of heart function (14, 15).

The heart is the most metabolically active organ in the body and has the highest content of mitochondria of any tissue (16). This is needed to meet the ever-demanding energy requirement for contraction and relaxation and about 95% of cellular ATP is utilized to support the contraction–relaxation cycle within the myocardium (16). Indeed, mitochondria contribute to the proper function of cardiomyocytes by multiple mechanisms (17). Beyond contraction, mitochondria also fine-tune cellular calcium homeostasis, apoptosis and oxidative stress and it has become increasingly clear that mitochondrial dysfunction is the source of heart energy deprivation in cardiomyopathies and heart failure of diverse etiologies (16).

Several myocardial mitochondrial enzymes were found to be selectively depressed in CCC as compared to other cardiomyopathies, including ATP synthase α , multiple fatty acid β -oxidation enzymes and creatine kinase activity [(18–20) and our unpublished observations] additionally, *in vivo* myocardial ATP production was shown to be reduced in CCC as determined by ³¹P-NMR spectroscopy (21). Taken together, results suggest that mitochondrial dysfunction may contribute to the worse prognosis of CCC. Since inflammatory cytokines IFN- γ and TNF- α are abundantly produced and are top upstream regulators of gene expression changes in CCC myocardium (9), and the sustained ROS and RNS production are known inducers of cardiomyocyte damage and mitochondrial dysfunction (22, 23), we hypothesized that such cytokines contribute to mitochondrial dysfunction observed in CCC (5).

Thus, this study aimed to investigate 1) whether CCC myocardial tissue display increased levels of mitochondrial dysfunction and oxidative/nitrosative stress markers compared to non-chagasic dilated cardiomyopathy biopsies (DCM); 2) whether IFN- γ and TNF- α promote mitochondrial dysfunction and nitro-oxidative stress in the human cardiomyocyte cell line AC-16; 3) whether the mitochondrial dysfunction in cardiomyocytes is

triggered by the cytokines through the STAT1/NF- κ B/NOS2 signaling pathway; 4) whether agonists of mitochondrial protection pathways, such as AMPK, NRF2 and SIRT1 ameliorate mitochondrial dysfunction driven by the IFN- γ and TNF- α .

METHODS

Ethics Statement

Tissue collection was approved by the Institutional Review Board of the University of São Paulo, School of Medicine (CAPPesq approval number 852/05) and written informed consent was obtained from the patients. All experimental methods comply with the Helsinki declaration.

Patients and Sample Collection

Human left ventricular free wall heart tissue was collected from end-stage heart failure CCC patients (n=40) and non-chagasic dilated cardiomyopathy (DCM, n=31) patients (NYHA class 3 and 4) at the moment of heart transplantation. CCC patients presented positive *T. cruzi* serology and typical heart conduction abnormalities (right bundle branch block and/or left anterior division hemiblock) and had a histopathological diagnosis of myocarditis, fibrosis and hypertrophy. All heart tissue samples were cleared from pericardium and fat and quickly frozen in liquid nitrogen and stored at -80°C. The list of patients is described in **Table 1**.

Reagents

Interferon-gamma (IFN- γ , 300-02) and tumor necrosis factor alpha (TNF- α , 300-01A) were purchased from Peprotech. AMPK agonists Metformin hydrochloride (PHR1084), AICAR (A9978), Resveratrol (R5010); SIRT1 agonist SRT1720 (SRT1720) and NRF-2/HO-1 agonist Protoporphyrin IX cobalt chloride (C1900) were purchased from Sigma-Aldrich. Inhibitors of NF- κ B Emodin (E7881) and JSH23 (J4455); IKK β inhibitor IKK16 (SML1138); STAT1 inhibitor Fludarabine (F2773) and NOS2 inhibitor 1400W (W4262) were purchased from Sigma-Aldrich. Inhibitors of MEK1 and MEK2 PD98059 (ab120234), JNK SP600125 (ab120065) and MAPK SB202190 (ab120638) were purchased from Abcam. Reagents were dissolved in DMSO or water (Metformin and AICAR) at 100x stock concentrations and maximum DMSO concentration used in cells was 1%. Monoclonal antibody against 3-nitrotyrosine (ab61392) was purchased from Abcam. Monoclonal antibody against NF- κ B p65 (sc-8008) was purchased from Santa Cruz Biotechnology.

Cell Culture

Human adult ventricular cardiomyocytes cell line AC-16 was kindly provided by Dr. Mercy Davidson (Columbia University, USA) (24). The cell line was propagated using Dulbecco's modified Eagle's medium/F-12 medium with 12.5% fetal bovine serum (FBS) without antibiotics for no longer than 8 passages. AC-16 cells were screened monthly for mycoplasma contamination (MycoAlert Mycoplasma Detection Kit, Lonza)

TABLE 1 | List of the patients included in this study.

| Code | Etiology | Sex | Age | EF (%) |
|---------|------------|-----|-----|--------|
| CCC-S01 | Severe CCC | F | 61 | 27 |
| CCC-S02 | Severe CCC | M | 60 | 35 |
| CCC-S03 | Severe CCC | F | 64 | 30 |
| CCC-S04 | Severe CCC | M | 61 | 21 |
| CCC-S05 | Severe CCC | F | 60 | 20 |
| CCC-S06 | Severe CCC | M | 41 | 19 |
| CCC-S07 | Severe CCC | M | 15 | 17 |
| CCC-S08 | Severe CCC | M | 28 | 29 |
| CCC-S09 | Severe CCC | M | 32 | 12 |
| CCC-S10 | Severe CCC | F | 32 | 19 |
| CCC-S11 | Severe CCC | M | 41 | 23 |
| CCC-S12 | Severe CCC | M | 58 | 20 |
| CCC-S13 | Severe CCC | M | 36 | 25 |
| CCC-S14 | Severe CCC | F | 47 | 27 |
| CCC-S15 | Severe CCC | F | 63 | 37 |
| CCC-S16 | Severe CCC | F | 44 | 25 |
| CCC-S17 | Severe CCC | M | 39 | 20 |
| CCC-S18 | Severe CCC | M | 58 | 25 |
| CCC-S19 | Severe CCC | M | 66 | 21 |
| CCC-S20 | Severe CCC | M | 50 | 25 |
| CCC-S21 | Severe CCC | M | 51 | 23 |
| CCC-S22 | Severe CCC | M | 58 | 29 |
| CCC-S23 | Severe CCC | M | 58 | 28 |
| CCC-S24 | Severe CCC | F | 66 | 25 |
| CCC-S25 | Severe CCC | F | 45 | 20 |
| CCC-S26 | Severe CCC | F | 60 | 24 |
| CCC-S27 | Severe CCC | F | 39 | 20 |
| CCC-S28 | Severe CCC | M | 51 | 35 |
| CCC-S29 | Severe CCC | F | 61 | 15 |
| CCC-S30 | Severe CCC | F | 47 | 35 |
| CCC-S31 | Severe CCC | F | 46 | 20 |
| CCC-S32 | Severe CCC | F | 61 | 27 |
| CCC-S33 | Severe CCC | F | 58 | 30 |
| CCC-S34 | Severe CCC | F | 49 | 15 |
| CCC-S35 | Severe CCC | M | 49 | 21 |
| CCC-S36 | Severe CCC | M | 62 | 21 |
| CCC-S37 | Severe CCC | M | 57 | 29 |
| CCC-S38 | Severe CCC | M | 59 | 17 |
| CCC-S39 | Severe CCC | M | 48 | 19 |
| CCC-S40 | Severe CCC | F | 54 | 36 |
| DCM-01 | DCM | M | 52 | 30 |
| DCM-02 | DCM | F | 32 | 20 |
| DCM-03 | DCM | F | 24 | 29 |
| DCM-04 | DCM | M | 46 | 25 |
| DCM-05 | DCM | M | 15 | 29 |
| DCM-06 | DCM | M | 55 | 25 |
| DCM-07 | DCM | F | 29 | 25 |
| DCM-08 | DCM | M | 36 | 14 |
| DCM-09 | DCM | M | 26 | 25 |
| DCM-10 | DCM | M | 42 | 9 |
| DCM-11 | DCM | | | |
| DCM-12 | DCM | F | 57 | 27 |
| DCM-13 | DCM | M | 48 | 39 |
| DCM-14 | DCM | F | 66 | 20 |
| DCM-15 | DCM | M | 43 | 26 |
| DCM-16 | DCM | M | 53 | 19 |
| DCM-17 | DCM | M | 39 | 28 |
| DCM-18 | DCM | M | 58 | 21 |
| DCM-19 | DCM | F | 12 | 22 |
| DCM-20 | DCM | F | 56 | 18 |
| DCM-21 | DCM | M | 29 | 26 |
| DCM-22 | DCM | M | 61 | 27 |
| DCM-23 | DCM | M | 15 | 20 |

(Continued)

TABLE 1 | Continued

| Code | Etiology | Sex | Age | EF (%) |
|--------|----------|-----|-----|--------|
| DCM-24 | DCM | F | 58 | 17 |
| DCM-25 | DCM | M | 28 | 16 |
| DCM-26 | DCM | M | 56 | 16 |
| DCM-27 | DCM | M | 27 | 25 |
| DCM-28 | DCM | F | 53 | 27 |
| DCM-29 | DCM | M | 51 | 15 |
| DCM-30 | DCM | M | 56 | 35 |
| DCM-31 | DCM | M | 37 | 16 |

and were authenticated by PCR-based DNA Profiling (Eurofins Genomics, France).

Cell Stimuli and Investigation of Molecular Pathways

AC-16 cells were seeded (10^3 cells per 0.34 cm^2) for 24 hours (h) prior to stimulation with 10 ng/ml IFN- γ , or with 5 ng/ml TNF- α or with IFN- γ plus TNF- α for 48h. In certain experiments, cells were treated with serially diluted, by a factor of two, with agonists of AMPK (metformin, AICAR and resveratrol), NRF2 (Protoporphyrin IX cobalt chloride), SIRT1 (SRT1720) or inhibitors of NF- κ B (Emodin, JSH23), IKK β (IKK16), STAT1 (Fludarabine), NOS2 (1400W), MEK1 and MEK2 (PD98059), JNK (SP600125) and MAPK (SB202190) in combination or not with IFN- γ , TNF- α or both. Compound doses that provided the highest effect on mitochondrial $\Delta\Psi$ m and less than 10% loss on cell number were selected for further analysis.

LDH Release Assay

The release of lactate dehydrogenase (LDH) in the conditioned media is an indicator of cell cytotoxicity. AC-16 cardiomyocytes were stimulated with the cytokines for 48h and the LDH content was quantified using a LDH-Cytotoxicity Assay Kit II (Abcam) using SpectraMax[®] Paradigm[®].

Mitochondrial Membrane Potential and Cellular ROS

AC-16 cardiomyocytes were seeded in 96 well black plate (Corning 3603) at a density of 10^3 cells per 0.34 cm^2 and incubated in a humidified incubator at 37°C and 5% CO_2 for 24h. Then, cells were stimulated with the cytokines with or without the agonists/inhibitors for 48h. For mitochondrial membrane potential, cells were multi-labeled with $1.0 \mu\text{M}$ of tetramethylrhodamine, methyl ester, perchlorate (TMRM, ThermoFisher Scientific), 400 nM of MitoTracker DeepRed (ThermoFisher Scientific), NucGreen dead (ReadyProbes) and $1.0 \mu\text{M}$ of Hoechst 33342 (ThermoFisher Scientific) at 37°C 5% CO_2 for 30 min. Cells were washed once with Hanks' Balanced Salt solution containing calcium and magnesium (HBSS++). Micrographs were captured using ImageXpress Micro XLS Widefield High-Content Analysis system (Molecular Devices) at 100x magnification and the mitochondrial membrane potential was evaluated in live cells (NucGreen negative) using the web-based software Columbus 2.7.1.133403 (PerkinElmer Inc.). Overlapped TMRM and Mitotracker DeepRed fluorescence

were considered for the measurement of mitochondrial $\Delta\Psi$ m. For cellular ROS, stimulated cells were washed once with HBSS++ and then labeled in the dark with $10 \mu\text{M}$ of H_2DCFDA (Abcam) and $1.0 \mu\text{M}$ of Hoechst 33342 (ThermoFisher Scientific) at 37°C 5% CO_2 for 30 min and fluorescence were measured at Ex/Em 485/535 nm and 360/461 nm respectively using SpectraMax[®] Paradigm[®].

NF- κ B Translocation

Stimulated AC-16 cells were fixed with 4% paraformaldehyde, pH=7.4 for 15 min at room temperature (RT), then incubated with ice-cold 0.5% Triton-X for 15 min and blocked with 3% BSA in PBS for 1h RT. Cells were incubated with 1:50 monoclonal mouse anti-NF- κ B p65 (SC-8008, Santa Cruz Biotechnology) in PBS + 1% BSA overnight at 4°C , then washed and kept in the dark with goat anti-mouse IgG1 Cross-Adsorbed Secondary Antibody (P-21129, ThermoFisher Scientific) 1:1000 in PBS + 1% BSA for 1h RT. Nuclei were stained with $2.5 \mu\text{g/ml}$ of DAPI (ThermoFisher Scientific) for 1h RT and images were acquired using ImageXpress Micro XLS Widefield High-Content Analysis system (Molecular Devices) at 40x magnification. Nuclear NF- κ B translocation was measured using the software Columbus 2.7.1.133403 (PerkinElmer Inc.).

Assessment of 3-Nitrotyrosine

Relative quantification of nitrated protein was performed by assessing 3-nitrotyrosine (3-NT) as a mean to detect nitrosative stress. Cardiac fragments and stimulated AC-16 cells were lysed in aqueous lysis solution containing 10 mM HEPES, 0.32 M sucrose, 0.1 mM Na_2EDTA , 1.0 mM dithiothreitol, $125 \mu\text{g/ml}$ PMSF and $1.0 \mu\text{l/ml}$ of proteinase inhibitor cocktail (Sigma), pH=7.4. Cardiac fragments were lysed in a tissue homogenizer (Precellys 24) pre-chilled 4°C using 2.8 mm ceramic beads for 3 cycles of 10s (seconds) with 15s intervals at 5500 rpm. Cells were lysed with 1.4 mm ceramic beads for 2 cycles of 30s with 10s intervals at 5000 rpm. Cardiac and cell lysates were clarified by centrifugation at $10,000 \text{ rcf}$ for 30 min at 4°C . A total of $5 \mu\text{g}$ of proteins was added to a nitrocellulose membrane, dried at 60°C for 15 min and then blocked with 3% Blotting Grade Blocker (BioRad) + TBS-T, for 1h at RT, under stirring. Membrane was washed with TBS-T and incubated with 1:1000 primary monoclonal antibody 3-nitrotyrosine (Abcam) in 3% BSA + TBS-T overnight 4°C . Then, membrane was washed twice and incubated with 1:1000 secondary antibody for 2h at RT, under agitation and protected from light. Fluorescence was captured using a scanner (LI-COR, Odyssey) or ECL revelation. Ponceau S images were captured using ImageQuant LAS-400 (GE Healthcare) and used for protein normalization.

ATP and Nitrite Production

Nitrite (NO_2^-) was measured in the conditioned media (phenol-free) of stimulated AC-16 cells using Griess Reagent Kit for Nitrite Determination (Molecular Probes) according to manufacturer's instructions. The cells were collected and lysed with TE buffer (100 mM Tris, 4 mM EDTA, pH=7.5) and ATP measured with a luciferase-based assay kit (ATP Determination Kit, ThermoFisher Scientific) according to manufacturer's instructions.

Nitrite absorbance and ATP-derived luminescence were measured in AC-16 cardiomyocytes using SpectraMax® Paradigm® (Molecular Devices). Measurement of nitrite in myocardial homogenates was performed by chemiluminescence with Nitric Oxide Analyzer (NOA 280, Zysense, United Kingdom).

Bioenergetic Function Analysis

A Seahorse XFe24 Analyzer (Agilent, Les Ulis, France) was used to survey bioenergetic function by measuring the oxygen consumption rate (OCR) and extracellular acidification rate (ECAR) of live cells. A quantity of 3,500 cells/0.32cm² were seeded in 150 µl of specific media, incubated for 1h at room temperature for adhesion and stimulated with 350 µl of specific media containing IFN-γ (10 ng/ml final) and TNF-α (5 ng/ml final) for 48h. OCR and ECAR were obtained from 90% confluent monolayer culture. All the experiments were done according to the protocol provided by the manufacturer. Briefly, the Agilent Seahorse Cell Mito Stress Test kit assesses mitochondrial function. Multiple parameters are obtained such as, basal respiration, ATP-linked respiration, maximal and reserve capacities, and non-mitochondrial respiration. The ATP Real-Time rate kit is the only assay that quantifies the ATP production from glycolysis and mitochondria simultaneously. The Agilent Seahorse Mito Fuel Flex Test kit measures the basal state of mitochondrial fuel oxidation in live cells by providing information on the relative contributions of glucose, glutamine and long-chain fatty acid oxidation to basal respiration. All analyses were performed using the software Wave 2.6.1 (Agilent, Les Ulis, France).

mtDNA Quantification

Total DNA from cardiac fragments and AC-16 cells was extracted using the DNA extraction kit FlexiGene DNA kit (QIAGEN) according to the manufacturer's instructions. PCR reaction was conducted using MAXIMA SYBR Green/ROX qPCR Master Mix (2x) (Thermo Scientific), 1 ng of template DNA and primers for mitochondrial encoded genes NADH: ubiquinone oxidoreductase subunit 1 (MT-ND1, 5 pmol/µl), cytochrome b (MT-CYTB, 5 pmol/µl), cytochrome c oxidase subunit 3 (MT-COXIII, 5 pmol/µl). Nuclear gene Telomerase (TERT, 15 pmol/µl) was used as the endogenous control. The PCR reactions were done in QuantStudio 12K (Applied Biosystem) with the following cycling program, 1 cycle of 50°C for 2 min, 95°C for 10 min, and 40 cycles of 50°C for 2 min, 95°C for 15s, and 60°C for 1 min. The ratio of mtDNA/nuDNA was calculated using the formula $2^{2\Delta\Delta CT}$ (25).

RNA Extraction and Sequencing

AC-16 cells were crushed with ceramic beads (VK05, diameter 1.4 mm) in 350 µl of RLT lysis buffer supplemented with 3.5 µl of β-mercapto-ethanol. Total RNA was extracted by using the RNeasy Mini Kit adapted with Trizol. RNA quantity was measured with a 2100 Bioanalyzer. Total RNA Sequencing was provided by Genewiz. Ribosomal RNAs have been depleted, and samples were prepared for sequencing according to the Illumina TruSeq RNA Preparation Kit and subjected to pairwise

sequencing (2x150 bp) with the Illumina Novaseq sequencer. After sequencing, the raw data files (fastq) were generated.

Differential Expression Analysis

Raw data quality has been verified with FastQC (v0.11.5). Low-quality reads, Illumina adapters and reads smaller than 20 nucleotides were removed with Trimmomatic (v0.39) (26), using default values for other options. Reads were aligned on GRCh37 (hg19) human reference genome from Ensembl using STAR (v2.5.4b) (27), specifying that the reads are 2x150 nucleotides (paired-end). Alignment quality has been checked by BAMQC (qualimap v2.2.1) (28). Gene quantification was done with featureCounts (v2.0.0) (29), using the exons as features and the genes as attributes. The DESeq2 package (v1.26.0) has been used to normalize data and to perform the differential expression test (30). To improve log₂ fold change accuracy, DESeq2 shrinkage function has been applied to the test results, and p-value was corrected with Benjamini-Hochberg method. Genes having an adjusted p-value ≤ 0.05 and a |log₂(FC)| ≥ 1.5 were considered to be differentially expressed genes (DEG). Then, results were analyzed by graphical methods with EnhancedVolcano (v1.4.0) and ggplot2 packages. Moreover, hierarchical clustering (HCA) based on spearman correlation was computed using gplots (v3.0.4) package, and principal component analysis (PCA) with factoextra (v1.0.7) and FactoMineR (v2.3) packages. We compared the AC-16 gene expression profile with the CCC gene expression profile (GSE84796) (31).

Proteomic Analysis of AC-16 Cardiomyocytes

The proteomic analysis of AC-16 cells under IFN-γ (10 ng/ml) and TNF-α (5 ng/ml) treatment was assessed by high-resolution mass spectrometric analysis. After 48h incubation with the combination of IFN-γ and TNF-α, proteins from AC-16 were extracted using lysis buffer (12 mM sodium deoxycholate, 12 mM *N*-lauroylsarcosine sodium salt, 100 mM Tris-HCl, pH=9.0) supplemented with protease inhibitor cocktail (Sigma-Aldrich). Total cell lysates from control and IFN-γ+TNF-α-treated cells were subject to in-solution trypsin digestion according to Phase Transfer Surfactant (PTS)-aided trypsin digestion protocol (32). Digested proteins were desalted using StageTip protocol (33) and dried by centrifugation under vacuum. Peptide samples were analyzed by RP-LC-MS/MS. For MS analysis, peptide mixtures were dissolved in 0.1% formic acid (solution A). Aliquots of 3 µl were automatically injected by a nano chromatography EASY-nLCII system (Thermo Scientific) on a 40 mm x 100 µm ID C-18 pre-column packed with Jupiter 10 µm resin (Phenomenex) and submitted to a chromatographic separation in a 100 mm x 75 µm ID C-18 column packed with Reprosil-Pur 3 µm C-18 beads (Dr. Maisch) coupled to an LTQ Orbitrap Velos (Thermo Scientific). Peptides were eluted with a linear gradient of 5-30% solution B (0.1% formic acid in acetonitrile) at 200 nl/min for 55 min. Spray voltage was set at 2.5kV and the mass spectrometer was operated in data dependent mode, in which one full MS scan was acquired in the m/z range of 400-1,800 at 60,000 resolution (at 400m/z)

followed by collisional induced fragmentation (CID) and MS/MS acquisition of the ten most intense ions from MS scan. Unassigned and +1 charge states were not subjected to fragmentation. The maximum injection time and AGC target were set to 100ms (milliseconds) and 1E6 for full MS, and 10ms and 1E4 for MS/MS. The minimum signal threshold to trigger fragmentation event, isolation window and normalized collision energy (NCE) were set to, respectively, 1000 cps, 2m/z and 35%. Dynamic peak exclusion (list size of 500) was applied to avoid the same m/z of being selected for the next 30s. Mass spectrometric raw data were processed using the software MaxQuant [version 1.6.2.10 (34)] and searched against a target human protein database (*Homo sapiens*, UniprotKB reviewed, 26,526 sequences) for protein identification and quantification. A False Discovery Rate (FDR) of 1% was required for both protein and peptide identifications. Label Free Quantification [MaxLFQ (35)] method was enabled. T-test was used for relative protein quantification and p-values were adjusted with the Benjamini-Hochberg method. Proteins with adjusted p-values ≤ 0.05 were considered to be differentially expressed.

Statistical Analysis

Cell data are reported as the ratio to non-treated cells. Cell viability was calculated as the ratio of the number of live cells (NucGreen-negative) and total cells (NucGreen-negative plus NucGreen-positive cells) $\times 100$. The results were expressed as mean \pm standard deviation (SD) or standard error of the mean (SEM) when noted. Statistical analysis was performed using the GraphPad Prism 9.2.0 software (GraphPad Software, Inc., CA). Data were tested for normality using the Shapiro-Wilk test before statistical tests. A nonparametric Mann-Whitney test was applied if data were not normally distributed. In all other cases, one-way ANOVA followed by Dunn's *post hoc* test was applied for multiple comparisons. A p-value < 0.05 was considered statistically significant.

RESULTS

Increased RNS and Decreased mtDNA Content in CCC Myocardial Tissue

We compared the nitro-oxidative profile of CCC and DCM myocardium. We first measured nitrite, a reactive nitrogen species (RNS) marker, by chemiluminescence with a NO Analyzer. We observed that CCC biopsies possess higher content of nitrite (132%; $p < 0.001$) in comparison to DCM lysates (Figure 1A). Detection of nitrotyrosine, an additional nitro-oxidative stress marker, was performed by dot-blot. We observed a significantly higher nitrotyrosine immunoreactivity in CCC (27%; $p < 0.01$) than DCM myocardial lysates (Figure 1B), indicating previous exposure to peroxynitrite. Also, mitochondrial DNA (MT-ND1) was lower in CCC (44%; $p < 0.001$) than DCM myocardial samples (Figure 1C) indicating a reduction of mitochondrial mass.

IFN- γ and TNF- α Impair AC-16 Cardiomyocyte Mitochondrial Membrane Potential ($\Delta\Psi_m$)

To investigate the role of IFN- γ and TNF- α on AC-16 cardiomyocytes, we stimulated the cells with several concentrations of IFN- γ and TNF- α and we measured the mitochondrial $\Delta\Psi_m$ in a high content screening platform. This was performed to establish the concentrations of the cytokines for subsequent analyses. We observed that IFN- γ impaired the $\Delta\Psi_m$ of AC-16 48h after stimulation and this impairment was enhanced when IFN- γ was combined with TNF- α ; TNF- α alone failed to cause statistically significant reductions in $\Delta\Psi_m$ (Figures 2A, B). We then selected the concentrations of 10 ng/ml of IFN- γ and 5 ng/ml of TNF- α and we detected that IFN- γ and TNF- α decrease $\Delta\Psi_m$ of mitochondria larger than $10 \mu\text{m}^2$ (Figures 2C, D) and the impact is higher on larger mitochondria (Supplementary Figure 1). We performed the LDH assay of the selected doses of IFN- γ (10 ng/ml) and TNF- α (5 ng/ml) and we observed no cytotoxicity (Figure 2E).

Cytokine-Stimulated AC-16 Cells Display Increased RNS and Decreased ATP and mtDNA

We investigated whether IFN- γ and TNF- α prompt nitrosative and oxidative stress in AC-16 cardiomyocytes. We used the Griess reaction to measure nitrite in the supernatant of 48h IFN- γ and TNF- α -stimulated AC-16 cells. Nitrite accumulation in the supernatant of AC-16 was increased by TNF- α alone and not by IFN- γ or their combination after 48h of incubation (Figure 3A). Detection of 3-NT was performed by dot-blot. We observed that although IFN- γ or the cytokine combination induced a median increase of 25-72% in protein nitration, this was not statistically significant, most likely due to the high dispersion (Figure 3B). Cytokine treatment of AC-16 cardiomyocytes with IFN- γ , TNF- α and combined for 48h reduced the amount of MT-ND1 as compared to non-stimulated cells (Figure 3C). We also measured reactive oxygen species (ROS) production in AC-16 stimulated cells stained with the fluorogenic dye H₂DCFDA and we found a 43% increase in ROS after stimulation with IFN- γ and TNF- α and 23% with TNF- α (Figure 3D). In addition, luminescence assay also showed that the TNF- α alone or combined with IFN- γ decreased 50% and 58%, respectively, the ATP content in AC-16 cardiomyocytes (Figure 3E).

Molecular Pathways Analysis of Cytokine-Induced $\Delta\Psi_m$ Reduction

In order to better understand the mechanisms by which IFN- γ and TNF- α diminish AC-16 $\Delta\Psi_m$, cells were treated with several compounds that activate or inhibit specific pathways. Each drug was titrated and the concentration of each compound was selected based on the highest restorative effect on $\Delta\Psi_m$ and no more than 10% loss on cell number (Supplementary Figures 2, 3). Concurrent treatment showed that agonists of signaling pathways related to response to stress and mitochondrial protection, such as AMPK (metformin, AICAR, resveratrol),

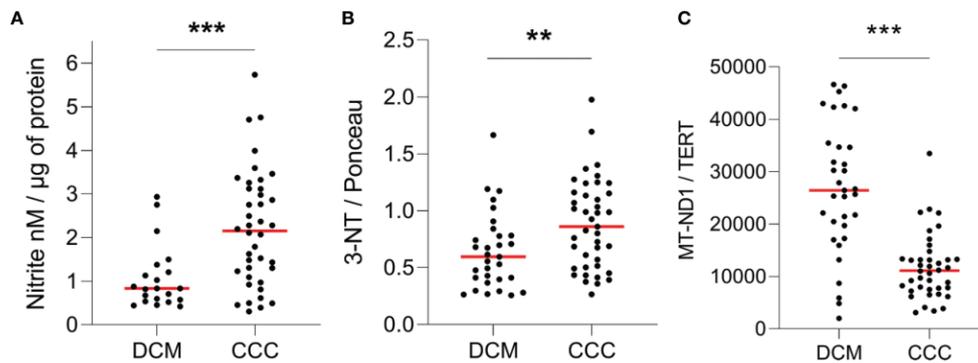


FIGURE 1 | Quantification of ROS, RNS and mitochondrial DNA in heart tissue. Lysates of left ventricular heart tissue were obtained from CCC (n = 40) and DCM (n = 31) patients and used for the quantification of **(A)** nitrite (nM) by chemiluminescence assay; **(B)** 3-nitrotyrosine by dot-blot and **(C)** copy number of the mitochondrial gene MT-ND1 by real time PCR. Red line: median; **p < 0.01; ***p < 0.001; Mann-Whitney test.

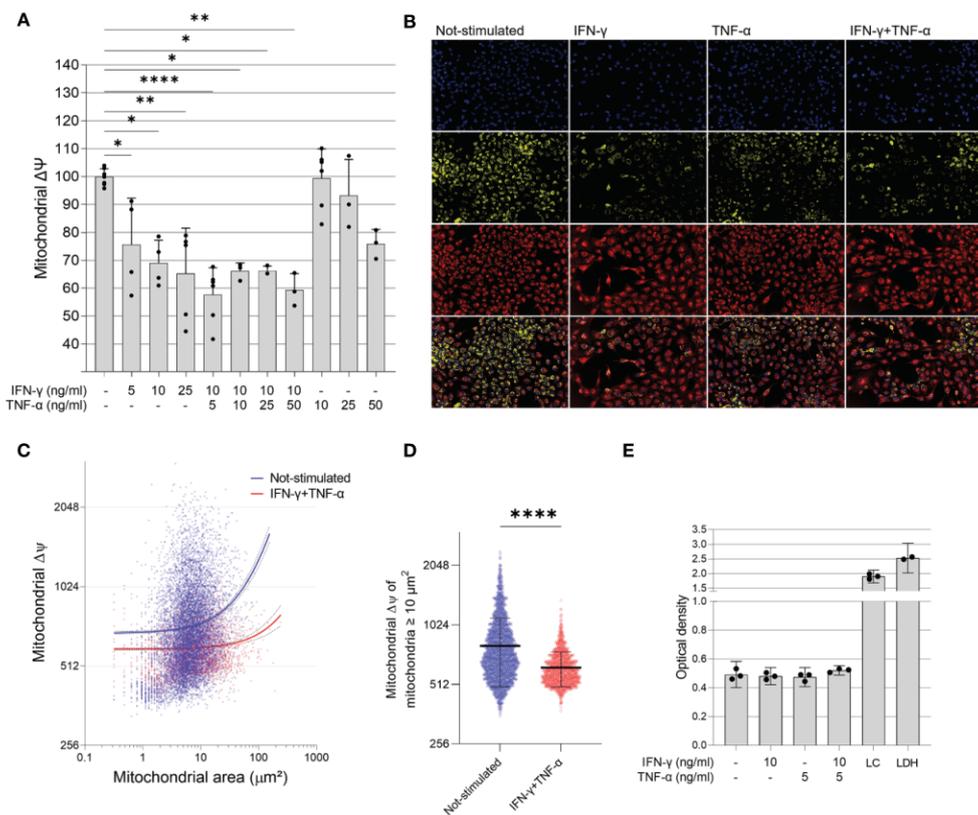


FIGURE 2 | Effect of IFN- γ and TNF- α on mitochondrial membrane potential in human cardiomyocytes AC-16 cells. Stimulated cells were multi-labelled with TMRM, Mitotracker DeepRed, NucGreen and DAPI. TMRM fluorescence was measured when colocalized with the fluorescence of Mitotracker Deep Red in live cells (NucGreen-negative). **(A)** IFN- γ and IFN- γ plus TNF- α downmodulate $\Delta\Psi_m$ 48h after stimulation. Data shown as percentage to not-stimulated cells. *p < 0.05; **p < 0.01; ****p < 0.0001; One-way Anova, Dunn's post test. **(B)** Representative micrographs of the effect of IFN- γ (10 ng/ml), TNF- α (5 ng/ml) and both on the fluorescence of TMRM. Magnification 100x. **(C)** Correlation between size (μm^2) and TMRM fluorescence intensity of segmented mitochondria of cells stimulated with 10 ng/ml of IFN- γ and 5 ng/ml of TNF- α . **(D)** $\Delta\Psi_m$ of mitochondria larger $\geq 10 \mu m^2$ of cells stimulated with 10 ng/ml of IFN- γ and 5 ng/ml of TNF- α . ****p < 0.0001 Mann-Whitney test. **(E)** Supernatant quantification of lactate dehydrogenase (LDH assay) on stimulated cells. Each dot in bar graphs represents an independent experimental replicate n \geq 3. LC: lysed AC-16 cells; LDH: lactate dehydrogenase positive control.

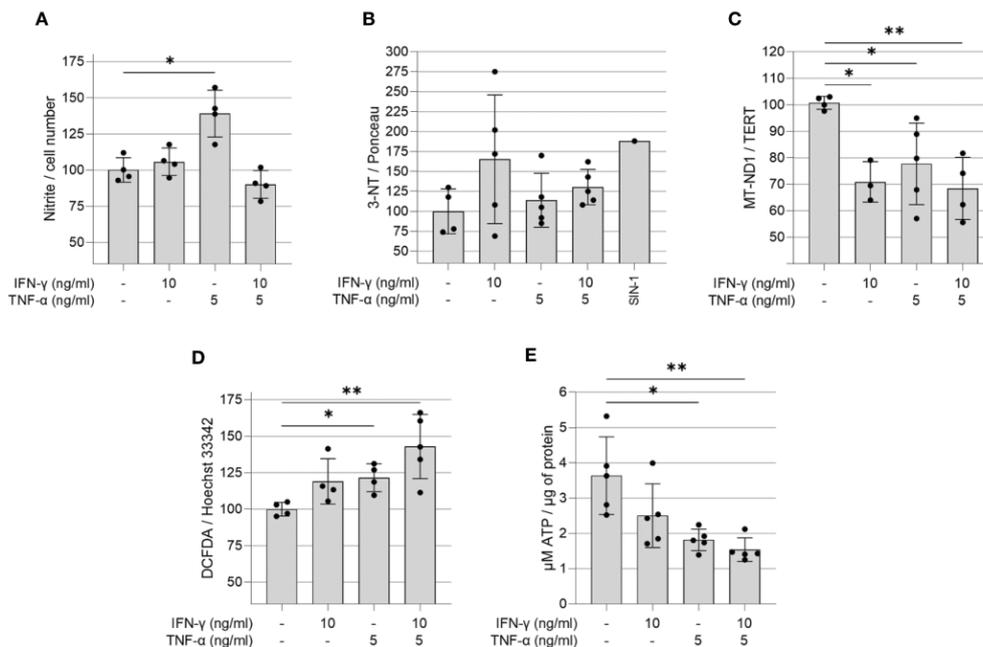


FIGURE 3 | Quantification of ROS, RNS, ATP and mitochondrial DNA in IFN- γ and TNF- α stimulated AC-16 cells. Human cardiomyocytes AC-16 cells were stimulated with 10 ng/ml of IFN- γ , 5 ng/ml of TNF- α or both for 48h. Then, cell lysates were used for the quantification of **(A)** nitrite (μ M) by Griess Reaction; **(B)** 3-nitrotyrosine by dot blot; **(C)** copy number of the mitochondrial gene MT-ND1 by real time PCR; **(D)** cellular reactive oxygen species (ROS) by fluorescence assay using the probe DCFDA and **(E)** ATP by luciferase-based assay. Data are shown as percentage to not-stimulated cells in **(A–D)** * $p < 0.05$; ** $p < 0.01$; one-way Anova Dunn's post test. Each dot represents an independent experiment $n \geq 3$.

NRF2 (Resveratrol and Protoporphyrin XI) and SIRT1 (SRT1720) restore $\Delta\Psi_m$ (Figures 4A–C). Additionally, inhibition of IFN- γ downstream molecules with selected antagonists such as fludarabine (STAT1 inhibitor), emodin (NF- κ B activation) JSH23 (NF- κ B nuclear translocation), IKK16 (I κ B kinase) and NOS2 (1400W) also restores AC-16 $\Delta\Psi_m$ (Figures 4D–G). Finally, inhibition of TNF- α -downstream molecules such as MEK (PD98059), JNK (SP600125) and MAPK (SB202190) also restores AC-16 $\Delta\Psi_m$ (Figures 4H–J) respectively). The effect of NF- κ B inhibitors on NF- κ B nuclear translocation was validated by additional experiments. Immunocytochemistry shows that IFN- γ and TNF- α promote NF- κ B nuclear translocation (Figure 4K), which was antagonized with emodin and JSH23 and simultaneously restored $\Delta\Psi_m$ in AC-16 cardiomyocytes (Figures 4L–M). In addition, Figure 4N shows that NOS inhibitor 1400W significantly decreased nitrite accumulation caused by TNF- α ($p < 0.001$). A schematic representation of the signaling pathways and the targets of the compounds are detailed in Supplementary Figure 4.

IFN- γ +TNF- α Costimulation on the AC-16 Cell Line Induces a Significant Decrease of Total, Glycolysis and Mitochondrial ATP Production

The effect of IFN- γ +TNF- α stimulation on AC-16 was investigated by measuring the two major production pathways in mammalian

cells (glycolysis and oxidative phosphorylation). Twenty-two independent measurements were done on non-stimulated AC-16 cells and 16 independent measurements were done on IFN- γ +TNF- α stimulated AC-16 cells. The OCR, ECAR and PER values were used to calculate glycolytic and mitochondrial ATP (Figures 5A–D). IFN- γ +TNF- α induced a significant decrease of total ATP production (33%; $p=0.0001$) (Figure 5E), glycolysis ATP production (30%; $p<0.0001$) (Figure 5F), and the mitochondrial ATP production (41%; $p=0.02$) (Figure 5G) was in line with the previous finding (Figure 3E). The percentage of oxidative phosphorylation was unchanged (Figure 5H).

IFN- γ +TNF- α Costimulation on the AC-16 Cell Line Induces an Increase of the Basal Respiration, Maximal Respiration, and Spare Respiratory Capacity

Metabolic flux analyses were performed on AC-16 cells with or without cytokine stimulation during 48h ($n=21$) (Figure 5I). Figure 5J illustrates the parameters analyzed. Cytokine-stimulated AC-16 cardiomyocytes increased basal respiration (53%; $p=0.0007$) compared to unstimulated cells ($n=20$) (Figure 5K). Also, the cytokines increased 54% ($p=0.0009$) the basal respiration used to drive ATP production (Figure 5L). Cytokine stimulation for 48h induces a proton leak (51%) indicating alteration of the mitochondria integrity ($p=0.021$) (Figure 5M). We also identified that cytokine stimulation

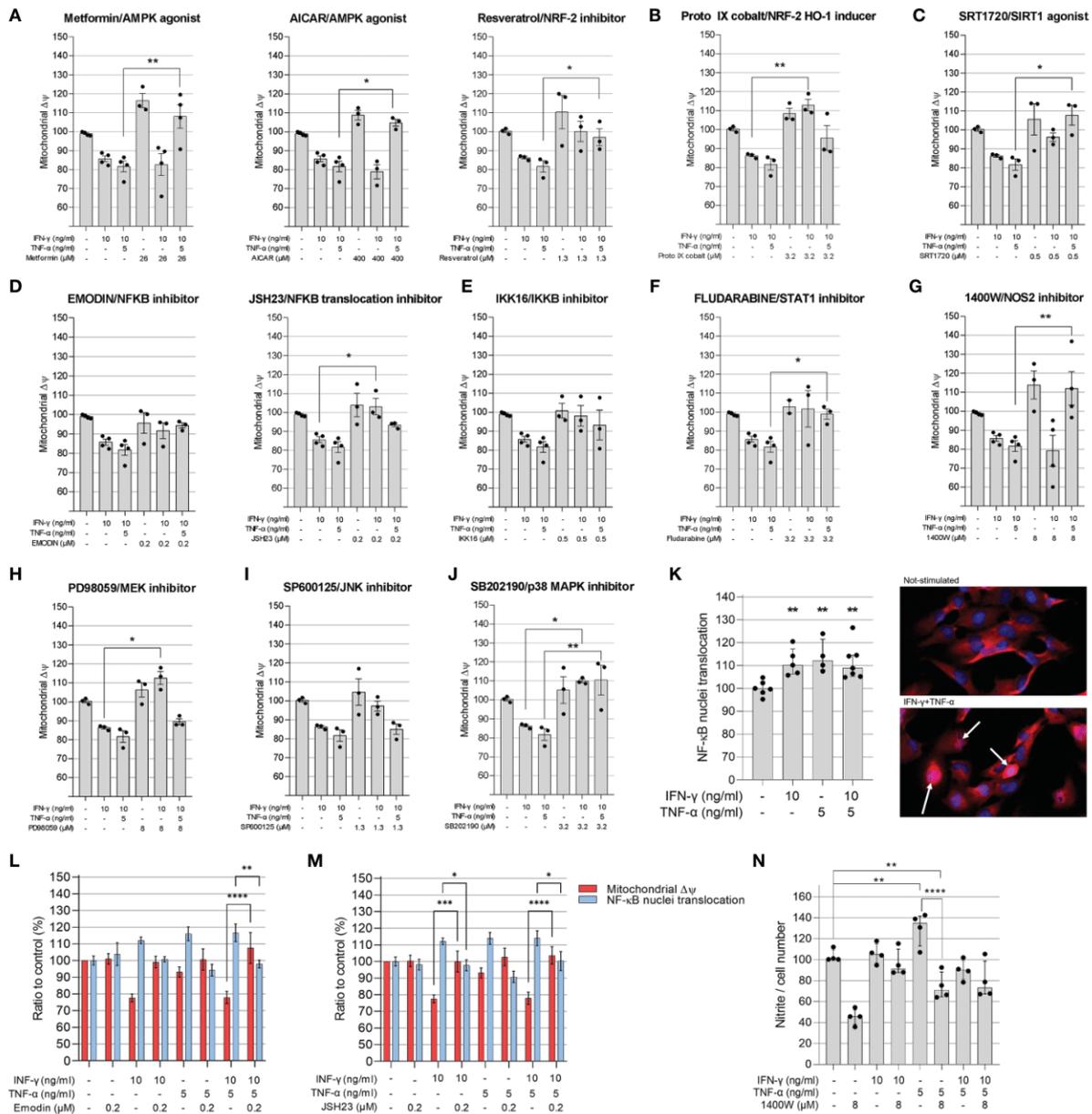
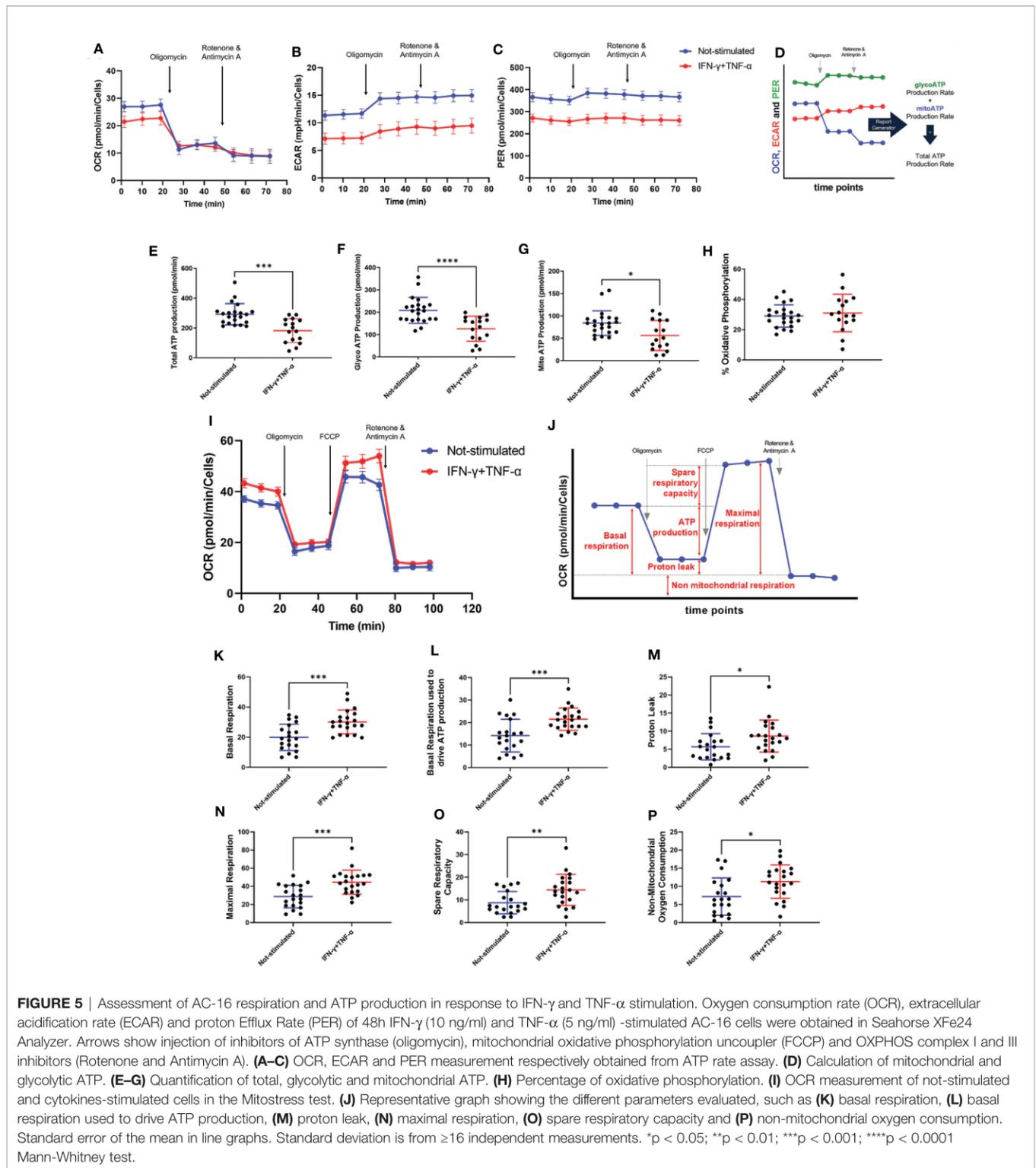


FIGURE 4 | Evaluation of compounds in the mitochondrial membrane potential of AC-16 cells. Selected doses of agonists of AMPK (A), NRF2 (B), SIRT1 (C) or inhibitors of NF-κB (D), IKKβ (E), STAT1 (F), NOS2 (G), MEK1 and MEK2 (H), JNK (I) and MAPK (J) were used alone or in combination with 10 ng/ml of IFN-γ or IFN-γ plus 5 ng/ml of TNF-α. Specific doses were selected based on the highest effect on mitochondrial ΔΨm and less than 10% loss on cell number. (K) The nuclear translocation of NF-κB in AC-16 cells was quantified by immunocytochemistry. Simultaneous measurement of ΔΨm and NF-κB translocation of cells stimulated with 10 ng/ml of IFN-γ or IFN-γ plus 5 ng/ml of TNF-α in combination with the NF-κB inhibitors (L) emodin and (M) JSH23. Statistics shown only when both ΔΨm and NF-κB translocation were significant. (N) the amount of nitrite (μM) in conditioned medium was measured by Griess reaction of cells stimulated with 10 ng/ml of IFN-γ or IFN-γ plus 5 ng/ml of TNF-α in combination with NOS2 inhibitor 1400W. Micrographs of NF-κB stains in the nucleus (white arrow). All data are shown as percentage to not-stimulated cells. Standard deviation is from ≥3 independent experiments. *p < 0.05; **p < 0.01; ****p < 0.0001 one-way Anova Dunn's post test.

increases the maximal respiration (44% higher) compared with unstimulated cells (p=0.0005) (Figure 5N). Similar results are obtained for the spare respiratory capacity, which is the difference between maximal respiration and basal respiration

(p=0.005) (Figure 5O). Cytokine stimulation also induced an increase in non-mitochondrial respiration (74%) compared to the non-mitochondrial respiration detected on unstimulated cells (p=0.014) (Figure 5P).



IFN- γ +TNF- α Costimulation Decreases AC-16 Cells' Dependency of Fatty Acid and Glutamine Oxidation

We investigated the effect of IFN- γ +TNF- α on AC-16 cell's dependency and capacity of fuel oxidation using Seahorse. We

measured oxygen consumption after injection of different inhibitors (combined or not), such as: UK5099, inhibitor of the mitochondrial pyruvate carrier (MPC, glucose oxidation), BPTES, inhibitor of glutaminase GLS1 (KGA, glutamine oxidation) and etomoxir, an inhibitor of carnitine

palmitoyltransferase-1 (CPT-1, fatty acid oxidation). We then calculated the dependency, which is the cell's reliance on a particular fuel pathway to maintain basal respiration, and capacity, i.e. the ability of mitochondria to oxidize a fuel when other fuel pathways are inhibited.

We found that 48h stimulation with IFN- γ +TNF- α decreased 62% ($p=0.0104$ $n=9$) the reliance of AC-16 cells on glutamine (Figures 6A, D) and 87% dependency of fatty acid oxidation ($p=0.0411$ $n\geq 9$) (Figures 6B, E), with no effect on glucose dependency (Figures 6C, F). In addition, cells' capacity to oxidize fatty acid and glucose increased 46% ($p=0.0037$ $n\geq 10$) (Figures 6H, K) and 105% ($p<0.0001$ $n\geq 11$) (Figures 6I, L) respectively after stimulation with the cytokines. No changes were detected in glutamine capacity (Figures 6G, J). A summary of the global fuel oxidation of not-stimulated and cytokine-stimulated cells is shown in Figure 6M.

Mitochondrial Proteins Are Differentially Expressed Both in CCC and Cytokine-Treated AC-16 Cardiomyocytes

Extracted proteins from AC-16 were submitted to LC-MS/MS and comparative quantitative analysis of 1,211 identified proteins resulted in 275 differentially expressed proteins (DEP) between IFN- γ +TNF- α -treated AC-16 and non-treated AC-16; 37 DEPs were mitochondrial proteins, 25 being down-modulated and 12 up-regulated (Supplementary Table 1). Among downregulated proteins, 6 belonged to the TCA cycle and OXPHOS (ACO2, DLAT, FH, CS, NDUFB2, NDUFB8). We found reduced expression of two mitochondrial protein import enzymes (TIMM23, TOMM22) two linked to protein synthesis (RARS, GRSF1) and 2 to maintenance of mitochondrial membrane function and polarization (TRAP1, IMMT) and reduced levels of the ion carriers SLC25A40, SLC25A11, SLC25A3. We found downregulated proteins belonging to the lipid beta oxidation (ACOT13, HADH, FASN, ACAT1) and involved in ATP metabolism, ATP synthase regulation and ATP transport to the cytoplasm (ATPIF1, USMG5, PPIF and SLC25A6). PRDX3 and PRDX5 were downregulated and SOD2 was upregulated, consistent with oxidative stress. Reactome analysis of differentially expressed mitochondrial proteins disclosed repressed mitochondrial beta-oxidation and TCA cycle, increased glycolysis, increased mitochondrial protein import, ketone body metabolism and detoxification of reactive oxygen species. IPA canonical pathways analysis indicated mitochondrial dysfunction, reduced sirtuin-1 signaling, reduced TCA cycle, downregulated ketolysis, upregulation of superoxide radical degradation. A comparison of DEPs both in the cytokine-stimulated AC-16 cardiomyocytes and CCC myocardium (data not shown; submitted for publication) identified 19 matching proteins. Twelve out of the 19 were mitochondrial or related to energy metabolism, and 8 of them were similarly modulated in both proteome datasets. Among matching and differentially expressed mitochondrial proteins, we observed reduced expression of proteins including HADH and ACAT1 (fatty acid beta-oxidation), NDUFB2 (OXPHOS), and FH (TCA), with increased expression of PKM (glycolysis) and LAP3 (apoptosis).

Mitochondrial Genes Are Differentially Expressed in Cytokine-Treated AC-16 Cardiomyocytes

The AC-16 cell line was stimulated with IFN- γ and TNF- α during 1h to 48h. Gene expression analysis was performed in the various time points taking as reference the T=0h time point. We considered differentially expressed genes, genes characterized by a corrected p -value <0.05 (Benjamini Hochberg) and absolute \log_2 Fold Change >1.5 . A total of 1443 DEGs were detected in all these comparisons (Supplementary Table 2). These DEGs include mainly protein coding genes ($n=1149$, 79.6%), as well as pseudogenes ($n=68$, 4.7%), antisense ($n=89$, 6.2%) and lncRNAs ($n=94$, 6.5%). The largest number of DEGs were detected at 6h or 12h (T=1h: $n=214$; T=6h: $n=858$; T=12h: $n=847$; T=24h: $n=588$; T=48h: $n=234$). Volcano plots and heatmaps are described in Supplementary Figure 5 and Supplementary Table 3. The main pathways are linked to the immune response, to inflammation, to intracellular defense mechanisms, to fibrosis and cardiac diseases. Just after 1h of cytokine stimulation AC-16 cells increased the expression of cytokines and chemokines such as CXCL1, CCL2, CXCL2 and IL8, IRF1 and IRF9, indicating an interferon and pro-inflammatory signaling cascade with an overexpression of TNFAIP3 and NFKBIA which act as NFKB inhibitors. After 6h of stimulation, a large number of cytokine and chemokine genes and their receptors was upregulated, including CCL2, CCL7, CCL8, CCL13, CX3CL1, CXCL1, CXCL2, CXCL3, CXCL5, CXCL9, CXCL10, CXCL11, IFNAR2, IFNGR2, STAT1, TNF, IL1A, IL1B, IL4R, IL6, IL7, IL7R, IL12A, IL15, IL15RA, IL20RB, IL21R, IL32, IL33 genes). We also performed the same experiments with IFN- γ or TNF- α alone (Figure 7 and Supplementary Table 4). Overall, we found 67 mitochondrial genes to be differentially expressed across at least one of the time points in response to combined cytokine stimulation. Differentially expressed mitochondrial genes peaked at 12h (46 genes) of which 25 were upregulated and 21 down-regulated (Supplementary Table 5). Among the top 10 canonical pathways of mitochondrial DEGs at 12h stimulus, we observed inhibition of the sirtuin signaling pathway, activation of aryl hydrocarbon receptor signaling, interferon signaling, and mitochondrial dysfunction. Toxicity-function pathways analysis at 12h indicated decreased transmembrane potential of mitochondria, fatty acid metabolism, mitochondrial dysfunction, pro-apoptosis, cardiac necrosis/cell death, cardiac hypertrophy among others (Supplementary Table 6). Mitochondrial dysfunction and decreased transmembrane potential were enriched toxicity functions in all time points starting at 6h (data not shown).

DISCUSSION

In this study, we investigated the mitochondrial function, nitro-oxidative profile, and gene and protein expression of myocardial samples from CCC patients. We also surveyed the effect of IFN- γ and TNF- α in AC-16 cardiomyocytes' mitochondria. These

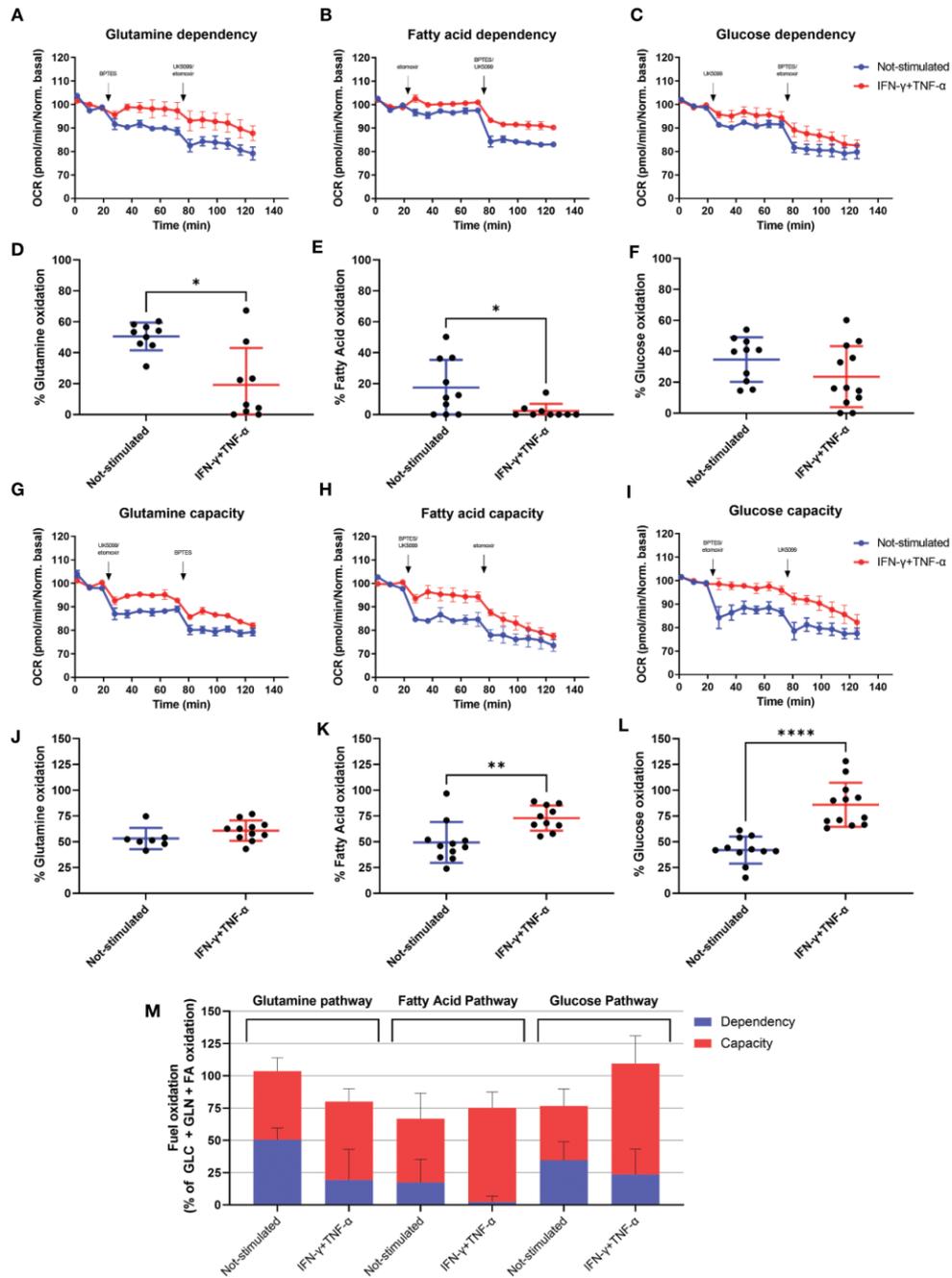


FIGURE 6 | Assessment of AC-16 fuel oxidation in response to IFN- γ and TNF- α stimulation. Oxygen consumption rate of 48h IFN- γ (10 ng/ml) and TNF- α (5 ng/ml)-stimulated AC-16 cells was obtained in Seahorse XFe24 Analyzer. Arrows show injection of inhibitors (combined or not) of molecules specific to different fuel oxidation pathways, such as; BPTES, inhibitor of glutaminase GLS1 (KGA, glutamine pathway), etomoxir, inhibitor of carnitine palmitoyltransferase-1 (CPT-1, fatty acid pathway) and UK5099, inhibitor of the mitochondrial pyruvate carrier (MPC, glucose pathway). **(A–F)** OCR dependency of glutamine, fatty acid oxidation and glucose. **(G–I)** OCR capacity of glutamine, fatty acid oxidation and glucose. **(J–L)** Global fuel oxidation. Standard error of the mean in line graphs. Standard deviation is from ≥ 9 independent measurements. * $p < 0.05$; ** $p < 0.01$; **** $p < 0.0001$ Mann-Whitney test.

cytokines are known to be selectively produced by the myocardium of CCC patients and not DCM (7–11, 36, 37). We have identified that CCC myocardium displays an increased nitro-oxidative stress profile, as well as reduced mtDNA content

in comparison to DCM samples and that these phenomena were also observed by IFN- γ /TNF- α treatment of AC-16 cardiomyocyte cell line. We performed mechanistic studies to better understand the role of the multiple signaling pathways in

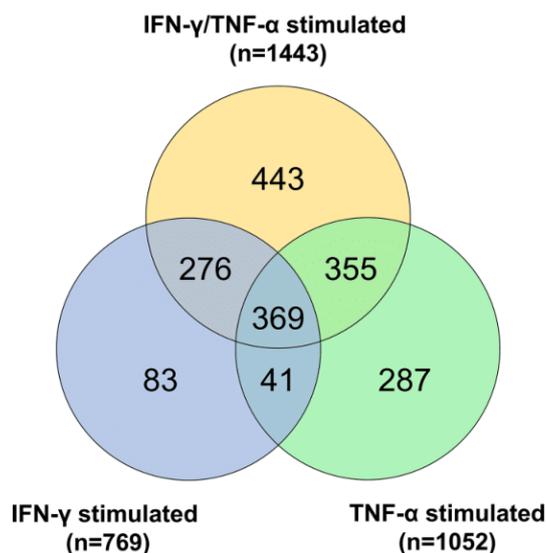


FIGURE 7 | Gene expression analysis on AC-16 cardiomyocytes. Cells were stimulated with IFN- γ (10 ng/ml) or TNF- α (5 ng/ml) or IFN- γ + TNF- α during 1 or 6 or 12 or 24 or 48h. On cardiomyocytes stimulated with TNF- α 1052 DEGs were detected whereas on cardiomyocytes stimulated with IFN- γ 769 DEGs were detected. Finally, on cardiomyocytes stimulated 48h with IFN- γ + TNF- α 1443 DEGs were detected. Venn diagram describes the DEGs shared by the various stimulations. Each stimulation was performed 4 times (4n).

the mitochondrial function of IFN- γ /TNF- α -treated AC-16 cardiomyocytes. We have found that the inhibition of STAT1/NF- κ B/NOS2 axis and activation of AMPK, NRF2 and SIRT1 signaling pathways promoted protective effects in the IFN- γ /TNF- α -induced impairment of mitochondrial $\Delta\Psi$. Pathways analysis of gene and protein expression involved mitochondrial dysfunction and decreased $\Delta\Psi$ m were found in cytokine-treated AC-16 cells. In addition, we observed a cytokine-induced reduction in ATP production at the expense of mitochondrial energy metabolism in AC-16 cardiomyocytes that paralleled that observed in CCC myocardial tissue.

IFN- γ , but not TNF- α alone, was shown to impair the mitochondrial $\Delta\Psi$ m of AC-16 cells. We also observed that concurrent treatment with IFN- γ and TNF- α results in a further decrease in $\Delta\Psi$ m, and this reduction is higher in larger mitochondria. This finding is especially crucial since the mitochondrial $\Delta\Psi$ m is the OXPHOS proton motive force that drives ATP production through the ATP synthesis complex and thus it is an essential mechanism for contraction and survival of cardiac cells (38). Several *in vitro* studies have found the suppressing effect of TNF- α (39–42) and IFN- γ plus TNF- α (43) in mitochondrial $\Delta\Psi$ m in different cell types.

IFN- γ exerts its deleterious effects in the mitochondria at least partially by potentiating TNF- α -mediated NF- κ B signaling (44). Activation of NF- κ B triggers transcriptional activity of NOS2, which in turn produces nitric oxide (NO) and in the presence of reactive oxygen species (ROS) produces peroxynitrite (ONOO⁻) one of the most toxic and highly oxidative reactive species with substantial effects in mitochondria (45, 46). In our observations, IFN- γ and/or TNF- α can increase NF- κ B nuclear translocation,

nitrite and ROS production, and 3-nitrotyrosine accumulation in AC-16 cardiomyocytes. Peroxynitrite and protein nitration cause inactivation of enzymes, poly(ADP-ribosylation), mitochondrial dysfunction, impaired stress signaling and also potentially inhibits myofibrillar creatine kinase (MM-CK), an important controller of heart contractility (47, 48) whose activity has been shown to be decreased in CCC (18). Increases in ROS by IFN- γ and/or TNF- α was also reported in several cell lines (39, 41, 49, 50), including cardiomyocytes (42, 51–53). Enhanced ROS production is associated with increased levels of lipid peroxidation, decreased mtDNA copy number and impaired OXPHOS capacity, affecting cardiomyocyte structure and function which triggers signaling pathways involved in myocardial remodeling and failure (54, 55). Enhanced oxidative stress has been observed in CCC heart tissue as measured by the accumulation of malondialdehyde (submitted for publication). We here reported reduction of mtDNA content in CCC myocardium and IFN- γ /TNF- α -stimulated cardiomyocytes. The mtDNA is a circular double stranded DNA located in the mitochondrial matrix and codes for 37 genes (56) and the 13 polypeptides are the essential subunits of the OXPHOS complexes I, III, IV, and V (56, 57). Deficiency in mtDNA replication was shown to cause ROS accumulation and oxidative stress in murine cardiomyocytes (58, 59), which is also indicative of a reduced mitochondrial mass. Thus, our data suggest that mtDNA reduction observed in the CCC myocardium might be linked to the oxidative stress observed in IFN- γ and TNF- α -treated human cardiomyocytes. This is corroborated by the gene and protein expression analysis in cytokine-treated AC-16 cardiomyocytes, where decreased levels of proteins involved in ATP generation mitochondrial protein

and ion import and mitochondrial structural maintenance proteins and upregulated expression of proteins involved in ATP catabolism and mitochondrial transition pore. In addition, we observed pathways analysis of CCC heart tissue indicative of mitochondrial dysfunction, increased oxidative stress, and cardiac necrosis, all pointing towards mitochondrial stress and reduced functional capacity. In line with these, stimulation with IFN- γ and TNF- α depleted ATP production in AC-16 cells. Previous studies described the dampening of energy metabolism enzymes in CCC heart tissue, such as ATP synthase α and creatine kinase activity in patients (18–20, 60) and studies with animal models correlated with this outcome (61). Decreased *in vivo* ATP production was also observed in the CCC myocardium (21). In addition, our group identified an accumulation of heterozygous pathogenic variations including loss-of-function and stopgain/truncation of key mitochondrial genes in CCC patients (62). The loss of function mutation in one of the studied families was dihydroorotate dehydrogenase (DHODH) R135C. DHODH is involved in the oxidative phosphorylation by donating electrons to complex III and treatment with DHODH inhibitor Brequinar in IFN- γ +TNF- α -treated AC-16 cardiomyocytes caused additive damage to mitochondrial $\Delta\Psi_m$ (62).

The inhibition of IFN- γ and TNF- α downstream molecules and pathways - STAT1/NF- κ B, NOS2 - was important for the restoration of AC-16 $\Delta\Psi_m$ and reduction of nitrite levels in AC-16 cardiomyocytes. The inflammatory milieu (IFN- γ , TNF- α , and IL-1b) enhanced ROS production in *T. cruzi* infected cardiomyocytes (51). Also, ROS production directly signaled the nuclear translocation of RelA (p65), NF- κ B activation in AC-16 cells (63), indicating a positive feedback loop of stimulation between oxidative stress and NF- κ B signaling. Studies reported that long-term sustained increase in ROS and RNS promotes cardiomyocyte dysfunction and apoptosis (64) resulting in reduction of mitochondrial $\Delta\Psi_m$, lipid beta-oxidation (65) and ATP generation (66, 67).

We found that the treatment of AC-16 cells with agonists of AMPK (resveratrol, AICAR and metformin) and NRF2 (protoporphyrin IX cobalt and resveratrol) rescued $\Delta\Psi_m$. AMPK and NRF2 are involved in the cellular response to oxidative stress by countering the damaging effects of NF- κ B and by promoting ATP production and regulation of important physiological processes to restore heart function, such as autophagy (68–72). Our data showed that these agonists ameliorate IFN- γ /TNF- α -damaging effect to cardiomyocytes $\Delta\Psi_m$. Similarly, activation of AMPK (metformin) was shown to inhibit the enhancing effect of IFN- γ on the DOX-induced cardiotoxicity and prolonged the survival time in DOX-treated mice (65). Indeed, a recent study showed that antioxidants such as resveratrol and mitochondria-targeted antioxidants have potential benefits for the control of oxidative stress in the myocardium of mice with experimental Chagas disease cardiomyopathy (73). Our work potentially found the mechanistic link for the findings that treatment of chronically *T. cruzi*-infected mice with SIRT1 and/or AMPK agonists SIRT1720, resveratrol and metformin or antioxidants reduced

myocardial NF- κ B transcriptional activity, inflammation and oxidative stress, resulting in beneficial results for restoration of cardiac function (73, 74).

Our transcriptomic profiling on cytokine-stimulated AC-16 cardiomyocytes over time (0 to 48h) showed that a cardiomyocyte can respond to inflammatory stimuli by producing inflammatory cell-attracting chemokines and inflammatory cytokines on its own as early as 1h after stimulus, perpetuating inflammation. Previous studies by our group in a subset of 10 out of the 30 CCC samples studied here observed increased mRNA expression of multiple chemokines and cytokines (9, 36, 75) in CCC heart tissue. At each time point, several genes associated with the mitogen-activated protein kinase (MAPK) signaling pathway were differentially expressed. This pathway is also one of the main inducers of the NF- κ B pathway that is activated after inflammatory stimulus, ischemia/reperfusion, in congestive heart failure, dilated cardiomyopathy, after ischemic and pharmacological preconditioning, and in hypertrophy of isolated cardiomyocytes (76). Pathway analyses of gene and protein expression in the IFN- γ and TNF- α stimulated AC-16 cardiomyocytes profile were consistent with disturbances of ATP production and increased levels of reactive oxygen species. These metabolic changes affect cardiac ion channel gating, electrical conduction, intracellular calcium handling, and fibrosis formation.

As shown in the functional energy metabolism experiments, the integrity of the mitochondria is affected (proton leak). Therefore, cytokine-treated cardiomyocytes significantly increased their respiration (basal respiration, maximal respiration, spare respiratory capacity, as well as the percentage of basal respiration used to produce ATP). However, as mitochondria are less numerous (as seen with the reduced mitochondrial DNA) and their integrity is affected, the production of mitochondrial and glycolytic ATP is decreased in association with a significant decrease of the glutamine and fatty acid oxidation dependency – a reduction in metabolic flexibility that is found in failing hearts (77). This is in line with the IFN- γ and TNF- α induced NF- κ B activation and activation of expression of NADP oxidases, and inducible nitric oxide synthase (NOS2), leading to the production of large amounts of NO and reactive nitrogen species (12, 78, 79) leading to synthesis of the peroxynitrite anion (ONOO⁻) production, a strong oxidant arising from the reaction of NO with superoxide radical (O₂⁻) (46).

This study has limitations, since we studied frozen human heart tissue samples and cytokine-treated human cardiomyocytes *in vitro*. Although some of the changes observed in IFN- γ and TNF- α -treated AC-16 cardiomyocytes closely parallel those observed in CCC heart tissue, this convergence is not proof that the findings in tissue are induced by the same cytokines, since several other mechanisms can induce nitro-oxidative stress and mitochondrial damage in CCC myocardium and as a consequence of heart failure. Conversely, our results with cytokine-treated cardiomyocytes can bring insight not only about Chagas disease, but also in other cardiac disorders where IFN- γ and TNF- α play a role, such as inflammatory cardiomyopathies of other etiology.

This study demonstrated that stimulation with IFN- γ and TNF- α in human cardiomyocytes causes mitochondrial damage, oxidative and nitrosative stress, paralleling events observed in the cytokine-rich CCC heart tissue. It is important to notice that the part of the CCC samples analyzed here have been previously assessed for cytokine expression and IFN- γ was among the most highly expressed cytokine, while IFN- γ and TNF- α were the top upstream regulators (9, 36). Both CCC myocardium and stimulated cells exhibited damaging profiles in markers of cellular stress and increased ROS and RNS, and decreased ATP. Several mitochondrial-related pathways important for mitochondrial integrity, function and ATP production were dysregulated in cytokine-stimulated cells. Also, cytokine-stimulated cells exhibited impaired $\Delta\Psi_m$ and increased ROS and RNS and higher

amounts of ROS. It is important to point out the direct involvement of the STAT1/NF- κ B/NOS2 signaling pathway in the damaging effects of IFN- γ and TNF- α in cardiomyocytes' $\Delta\Psi_m$, as well as the restorative effects of stimulating the AMPK, NRF2 and SIRT1 pathways (Figure 8). Our results suggest that cytokine-induced disturbances in mitochondrial function and energy metabolism might play a role in the worse prognosis of Chagas disease cardiomyopathy. Therapy targeting mitochondrial function and energy imbalance should thus in principle be beneficial to restore cardiac function in CCC and other IFN- γ -dependent inflammatory heart diseases, like viral myocarditis and inflammatory cardiomyopathy of other etiologies, age-related myocardial inflammation and functional decline (80), myocardial infarction (81), and anthracycline antitumoral agent cardiotoxicity (65).

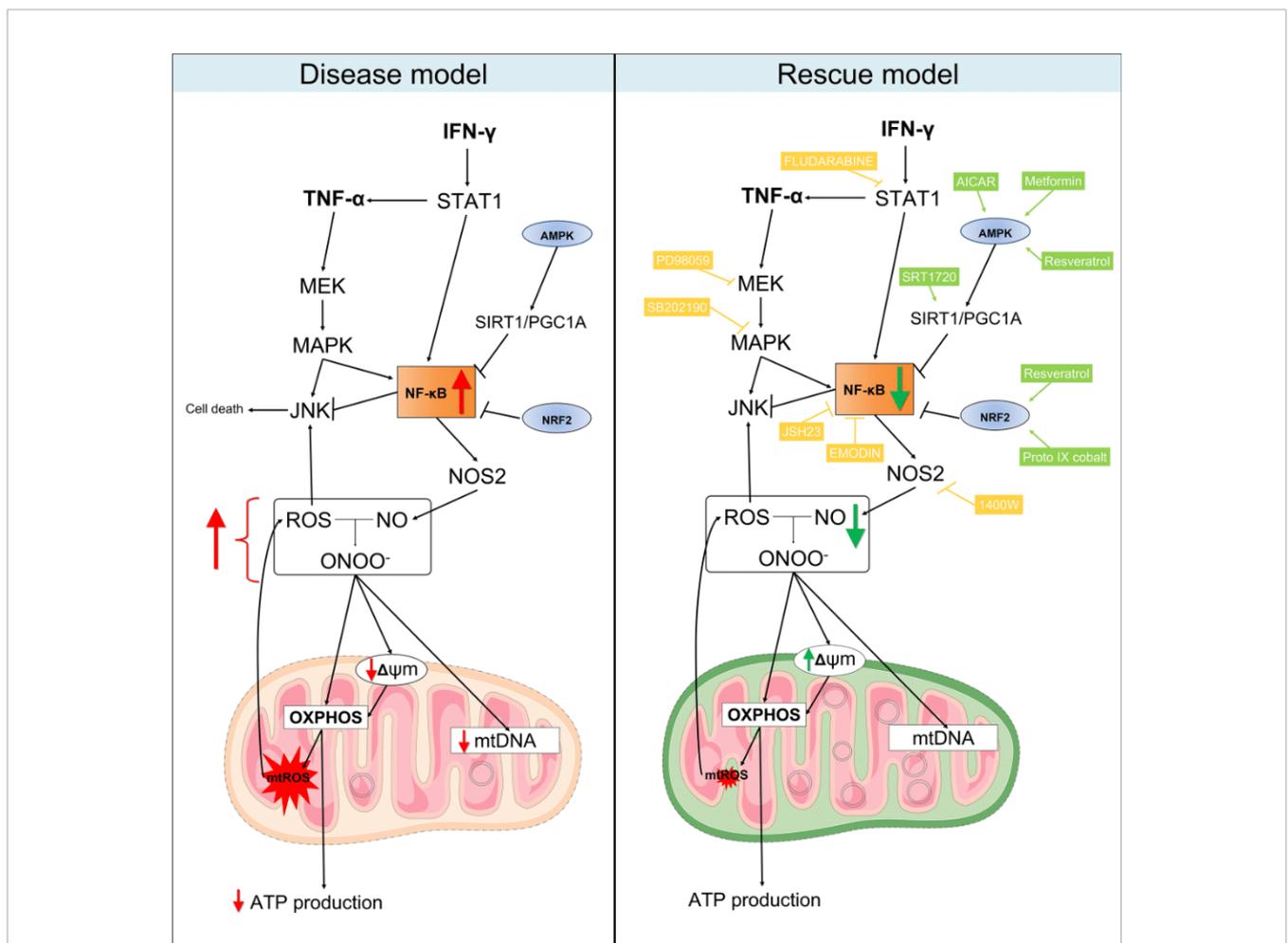


FIGURE 8 | Proposed model for the mitochondrial dysfunction in Chagas disease cardiomyopathy. Left panel: in summary, our results showed that heart tissues from CCC patients have increased production of RNS and reduced content of mitochondrial DNA in comparison with patients with DCM. Similarly, stimulation of AC-16 cardiomyocytes with IFN- γ /TNF- α increased ROS, RNS and proton leak, impaired $\Delta\Psi_m$, depleted ATP production and changed the metabolic profile of the cells. Right panel: in our rescue model, we showed that pharmacological inhibition of molecules involved in the IFN- γ /TNF- α /NF- κ B/NOS2 pathway ameliorated the $\Delta\Psi_m$ and NO production. Additionally, activation of AMPK/SIRT1 and NRF2 had beneficial impact on the $\Delta\Psi_m$. Thus, we hypothesize that mitochondrial dysfunction is driven by the excessive production of IFN- γ /TNF- α in CCC myocardium and is an essential component for the poor prognosis of Chagas disease cardiomyopathy. Mitochondria-targeted therapies might improve CCC disease progression. Compounds in yellow are inhibitors; Compounds in light green are agonists. Connecting arrows indicate activation and flat line means inhibitory interaction. Red and green arrows indicate the changes observed before and after treatment with the compounds.

Further studies with induced pluripotent stem cells derived cardiomyocytes (iPS-CM) can be employed to investigate, in a personalized, patient-specific manner, the effect of the cytokines in mitochondrial function.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding authors.

ETHICS STATEMENT

The protocol was also approved by the INSERM Internal Review Board and the Brazilian National Ethics in Research Commission (CONEP). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

Study design: EC-N, CC, JN, FL, and JK. Phenotype characterization: EB, RS, FB, and PP. Experimental analysis: JN, PA, RA, EK, AH, LI, DA-S, KS, RV, DL, SB, FG, MT, and PT. Statistical analysis: HN, JN, PB, DG, and MM. Manuscript preparation: EC-N, CC, JN, and VP. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Institut National de la Santé et de la Recherche Médicale (INSERM); the Aix-Marseille University (grant number: AMIDEX “International_2018” MITOMUTCHAGAS); the French Agency for Research (Agence Nationale de la Recherche-ANR (grant numbers: “Br-Fr-Chagas”, “landscardio”); the CNPq (Brazilian Council for Scientific and Technological Development); and the FAPESP (São Paulo State Research Funding Agency Brazil (grant numbers: 2013/50302-3, 2014/50890-5); the National Institutes of Health/USA (grant numbers: 2P50AI098461-02 and 2U19AI098461-06). This work was founded by the Inserm Cross-Cutting Project GOLD. This project has received funding from the Excellence Initiative of Aix-Marseille University - A*Midex a French “Investissements d’Avenir programme”- Institute MarMaRa AMX-19-IET-007. JN was a recipient of a MarMaRa fellowship. ECN and JK are recipients of productivity awards by CNPq. The funders did not play any role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

ACKNOWLEDGMENTS

We thank Andréia Kuramoto Takara (Laboratory of Immunology, InCor) and Victor Debbas (Laboratory of Vascular Biology, InCor) for their always helpful technical assistance.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2021.755862/full#supplementary-material>

Supplementary Figure 1 | Stratification of mitochondria area and membrane potential after stimulation with IFN- γ and TNF- α in human cardiomyocytes AC-16 cells. Cells were stimulated with 10 ng/ml of IFN- γ + 5 ng/ml of TNF- α for 48 hours and then multi-labelled with TMRM, Mitotracker DeepRed, NucGreen and DAPI. TMRM fluorescence was measured when colocalized with the fluorescence of Mitotracker Deep Red in live cells (NucGreen-negative). Mitochondria were stratified based on their area in: lower than 5 (**A**), between 5-10 (**B**), between 10-25 (**C**), between 25-50 (**D**) and (**E**) higher than 50 μm^2 . (**F**) average (\pm SEM) value obtained from the mitochondrial membrane potential of each stratification of mitochondrial size. **** $p < 0.0001$; Mann Whitney test.

Supplementary Figure 2 | Compound screening in AC-16 mitochondrial membrane potential. AC-16 cells were treated with serially diluted agonists of AMPK (**A**), NRF2 (**B**), SIRT1 (**C**) or inhibitors of NF- κ B (**D**), IKK β (**E**), STAT1 (**F**), NOS2 (**G**), MEK1 and MEK2 (**H**), JNK (**I**) and MAPK (**J**) alone (blue line) or in combination with 10 ng/ml IFN- γ (red line) or 10 ng/ml of IFN- γ plus 5 ng/ml TNF- α (green line) for 48 hours. Specific doses were selected based on the highest effect on mitochondrial $\Delta\Psi\text{m}$ and less than 10% loss on cell number. The first dot of the titration (to the left) is without compounds. Grey arrows indicate concentrations selected for. All data are shown as percentage to not-stimulated cells. Standard error of the mean is from 3 independent experiments.

Supplementary Figure 3 | Compound screening in the cell number of AC-16. AC-16 cells were treated with serially diluted agonists of AMPK (**A**), NRF2 (**B**), SIRT1 (**C**) or inhibitors of NF- κ B (**D**), IKK β (**E**), STAT1 (**F**), NOS2 (**G**), MEK1 and MEK2 (**H**), JNK (**I**), MAPK (**J**) and DMSO (**K**) alone (blue line) or in combination with 10 ng/ml of IFN- γ (red line) or 10 ng/ml of IFN- γ plus 5 ng/ml of TNF- α (green line) for 48 hours. The first dot of each titration (to the left) is without compounds. Grey arrows indicate concentrations selected for. Standard error of the mean is from 3 independent experiments.

Supplementary Figure 4 | Schematic representation of the inhibitors and agonists and their targets in the IFN- γ and TNF- α signaling pathways. The pro-inflammatory cytokine IFN- γ interacts with its transmembrane receptor IFNGR1 and signals mainly through the signal transducer and activator of transcription 1 (STAT1) and Interferon Regulatory Factor 1 (IRF1) intracellular transduction pathway to achieve transcriptional activation of IFN- γ -inducible genes, such as TNF- α and Nuclear Factor Kappa B (NF- κ B). Activation of NF- κ B leads to an upregulation of Nitric oxide synthase 2 (NOS2) gene, which in turn increases the amount of intracellular nitric oxide (NO), reacting to reactive oxygen species (ROS) to generate peroxynitrite (ONOO $^-$). Arrows indicate activation and flat line means inhibitory interaction. While ONOO $^-$ is important for pathogenic response, it causes damages to mitochondria, leading to reduction of mitochondria membrane potential and fragmentation. Proteins such AMPK and NRF2 are involved in the cellular response to oxidative stress countering the damaging effects of NF- κ B. Compounds highlighted in yellow are inhibitors and green are agonists.

Supplementary Figure 5 | Transcriptomic analysis on AC-16 cardiomyocyte cell line stimulated with IFN- γ and TNF- α . AC-16 cells were stimulated with IFN- γ and TNF- α during 1h to 48h. Gene expression analysis was done between the various time points taking as reference the T=0h time point. At each time is provided the volcano plot and the heatmap. Each stimulation was performed 4 times.

REFERENCES

- Savarese G, Lund LH. Global Public Health Burden of Heart Failure. *Card Fail Rev* (2017) 3(1):7–11. doi: 10.15420/cfr.2016:25:2
- Bocchi EA, Bestetti RB, Scanavacca MI, Cunha Neto E, Issa VS. Chronic Chagas Heart Disease Management: From Etiology to Cardiomyopathy Treatment. *J Am Coll Cardiol* (2017) 70(12):1510–24. doi: 10.1016/j.jacc.2017.08.004
- Issa VS, Ayub-Ferreira SM, Schroyens M, Chizzola PR, Soares PR, Lage SHG, et al. The Course of Patients With Chagas Heart Disease During Episodes of Decompensated Heart Failure. *ESC Heart Fail* (2021) p. 1460–71. doi: 10.1002/ehf2.13232
- Cunha-Neto E, Chevillard C. Chagas Disease Cardiomyopathy: Immunopathology and Genetics. *Mediators Inflamm* (2014) 2014:683230. doi: 10.1155/2014/683230
- Chevillard C, Nunes JPS, Frade AF, Almeida RR, Pandey RP, Nascimento MS, et al. Disease Tolerance and Pathogen Resistance Genes May Underlie. *Front Immunol* (2018) 9:2791. doi: 10.3389/fimmu.2018.02791
- Frade-Barros AF, Ianni BM, Cabantous S, Pissetti CW, Saba B, Lin-Wang HT, et al. Polymorphisms in Genes Affecting Interferon- γ Production and Th1 T Cell Differentiation Are Associated With Progression to Chagas Disease Cardiomyopathy. *Front Immunol* (2020) 11:1386. doi: 10.3389/fimmu.2020.01386
- Reis MM, Higuchi MEL, Benvenuti LA, Aiello VD, Gutierrez PS, Bellotti G, et al. An *in Situ* Quantitative Immunohistochemical Study of Cytokines and IL-2R α in Chronic Human Chagasic Myocarditis: Correlation With the Presence of Myocardial Trypanosoma Cruzi Antigens. *Clin Immunol Immunopathol* (1997) 83(2):165–72. doi: 10.1006/clin.1997.4335
- Abel LC, Rizzo LV, Ianni B, Albuquerque F, Bacal F, Carrara D, et al. Chronic Chagas' Disease Cardiomyopathy Patients Display an Increased IFN- γ Response to Trypanosoma Cruzi Infection. *J Autoimmun* (2001) 17(1):99–107. doi: 10.1006/jaut.2001.0523
- Laugier L, Ferreira LRP, Ferreira FM, Cabantous S, Frade AF, Nunes JP, et al. miRNAs may Play a Major Role in the Control of Gene Expression in Key Pathobiological Processes in Chagas Disease Cardiomyopathy. *PLoS Negl Trop Dis* (2020) 14(12):e0008889. doi: 10.1371/journal.pntd.0008889
- Reis DD, Jones EM, Tostes S, Lopes ER, Gazzinelli G, Colley DG, et al. Characterization of Inflammatory Infiltrates in Chronic Chagasic Myocardial Lesions: Presence of Tumor Necrosis Factor- α Cells and Dominance of Granzyme A $^{+}$, CD8 $^{+}$ Lymphocytes. *Am J Trop Med Hyg* (1993) 48(5):637–44. doi: 10.4269/ajtmh.1993.48.637
- Rocha Rodrigues DB, dos Reis MA, Romano A, Pereira SA, Teixeira VEP, Tostes S, et al. *In Situ* Expression of Regulatory Cytokines by Heart Inflammatory Cells in Chagas' Disease Patients With Heart Failure. *Clin Dev Immunol* (2012) 2012:361730. doi: 10.1155/2012/361730
- Hölscher C, Köhler G, Müller U, Mossmann H, Schaub GA, Brombacher F. Defective Nitric Oxide Effector Functions Lead to Extreme Susceptibility of Trypanosoma Cruzi-Infected Mice Deficient in Gamma Interferon Receptor or Inducible Nitric Oxide Synthase. *Infect Immun* (1998) 66(3):1208–15. doi: 10.1128/IAI.66.3.1208-1215.1998
- Wu Y, Antony S, Juhasz A, Lu J, Ge Y, Jiang G, et al. Up-Regulation and Sustained Activation of Stat1 Are Essential for Interferon-Gamma (IFN- γ)-Induced Dual Oxidase 2 (Duox2) and Dual Oxidase A2 (DuoxA2) Expression in Human Pancreatic Cancer Cell Lines. *J Biol Chem* (2011) 286(14):12245–56. doi: 10.1074/jbc.M110.191031
- Kuroda J, Sadoshima J. NADPH Oxidase and Cardiac Failure. *J Cardiovasc Transl Res* (2010) 3(4):314–20. doi: 10.1007/s12265-010-9184-8
- Sirker A, Zhang M, Murdoch C, Shah AM. Involvement of NADPH Oxidases in Cardiac Remodelling and Heart Failure. *Am J Nephrol* (2007) 27(6):649–60. doi: 10.1159/000109148
- Brown DA, Perry JB, Allen ME, Sabbah HN, Stauffer BL, Shaikh SR, et al. Expert Consensus Document: Mitochondrial Function as a Therapeutic Target in Heart Failure. *Nat Rev Cardiol* (2017) 14(4):238–50. doi: 10.1038/nrcardio.2016.203
- Chen L, Knowlton AA. Mitochondria and Heart Failure: New Insights Into an Energetic Problem. *Minerva Cardioangiol* (2010) 58(2):213–29.
- Teixeira PC, Santos RH, Fiorelli AI, Bilate AM, Benvenuti LA, Stolf NA, et al. Selective Decrease of Components of the Creatine Kinase System and ATP Synthase Complex in Chronic Chagas Disease Cardiomyopathy. *PLoS Negl Trop Dis* (2011) 5(6):e1205. doi: 10.1371/journal.pntd.0001205
- Wen JJ, Dhiman M, Whorton EB, Garg NJ. Tissue-Specific Oxidative Imbalance and Mitochondrial Dysfunction During Trypanosoma Cruzi Infection in Mice. *Microbes Infect* (2008) 10(10-11):1201–9. doi: 10.1016/j.micinf.2008.06.013
- Wan X, Gupta S, Zago MP, Davidson MM, Dousset P, Amoroso A, et al. Defects of mtDNA Replication Impaired Mitochondrial Biogenesis During Trypanosoma Cruzi Infection in Human Cardiomyocytes and Chagasic Patients: The Role of Nrf1/2 and Antioxidant Response. *J Am Heart Assoc* (2012) 1(6):e003855. doi: 10.1161/JAHA.112.003855
- Leme AM, Salemi VM, Parga JR, Ianni BM, Mady C, Weiss RG, et al. Evaluation of the Metabolism of High Energy Phosphates in Patients With Chagas' Disease. *Arq Bras Cardiol* (2010) 95(2):264–70. doi: 10.1590/S0066-782X2010005000099
- Lee HJ, Oh YK, Rhee M, Lim JY, Hwang JY, Park YS, et al. The Role of STAT1/IRF-1 on Synergistic ROS Production and Loss of Mitochondrial Transmembrane Potential During Hepatic Cell Death Induced by LPS/d-GalN. *J Mol Biol* (2007) 369(4):967–84. doi: 10.1016/j.jmb.2007.03.072
- Meyer A, Laverny G, Allenbach Y, Grelet E, Ueberschlager V, Echaniz-Laguna A, et al. IFN- β -Induced Reactive Oxygen Species and Mitochondrial Damage Contribute to Muscle Impairment and Inflammation Maintenance in Dermatomyositis. *Acta Neuropathol* (2017) 134(4):655–66. doi: 10.1007/s00401-017-1731-9
- Davidson MM, Nesti C, Palenzuela L, Walker WF, Hernandez E, Protas L, et al. Novel Cell Lines Derived From Adult Human Ventricular Cardiomyocytes. *J Mol Cell Cardiol* (2005) 39(1):133–47. doi: 10.1016/j.yjmcc.2005.03.003
- Quiros PM, Goyal A, Jha P, Auwerx J. Analysis of mtDNA/nDNA Ratio in Mice. *Curr Protoc Mouse Biol* (2017) 7(1):47–54. doi: 10.1002/cpmo.21
- Bolger AM, Lohse M, Usadel B. Trimmomatic: A Flexible Trimmer for Illumina Sequence Data. *Bioinformatics* (2014) 30(15):2114–20. doi: 10.1093/bioinformatics/btu170
- Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: Ultrafast Universal RNA-Seq Aligner. *Bioinformatics* (2013) 29(1):15–21. doi: 10.1093/bioinformatics/bts635
- Okonechnikov K, Conesa A, García-Alcalde F. Qualimap 2: Advanced Multi-Sample Quality Control for High-Throughput Sequencing Data. *Bioinformatics* (2016) 32(2):292–4. doi: 10.1093/bioinformatics/btv566
- Liao Y, Smyth GK, Shi W. Featurecounts: An Efficient General Purpose Program for Assigning Sequence Reads to Genomic Features. *Bioinformatics* (2014) 30(7):923–30. doi: 10.1093/bioinformatics/btt656
- Love MI, Huber W, Anders S. Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data With Deseq2. *Genome Biol* (2014) 15(12):550. doi: 10.1186/s13059-014-0550-8
- Laugier L, Frade AF, Ferreira FM, Baron MA, Teixeira PC, Cabantous S, et al. Whole Genome Cardiac DNA Methylation Fingerprint and Gene Expression Analysis Provide New Insights in the Pathogenesis of Chronic Chagas Disease Cardiomyopathy. *Clin Infect Dis* (2017) p. 1103–11. doi: 10.1093/cid/cix506
- Masuda T, Tomita M, Ishihama Y. Phase Transfer Surfactant-Aided Trypsin Digestion for Membrane Proteome Analysis. *J Proteome Res* (2008) 7(2):731–40. doi: 10.1021/pr700658q
- Rappsilber J, Mann M, Ishihama Y. Protocol for Micro-Purification, Enrichment, Pre-Fractionation and Storage of Peptides for Proteomics Using StageTips. *Nat Protoc* (2007) 2(8):1896–906. doi: 10.1038/nprot.2007.261
- Cox J, Mann M. MaxQuant Enables High Peptide Identification Rates, Individualized P.P.B.-Range Mass Accuracies and Proteome-Wide Protein Quantification. *Nat Biotechnol* (2008) 26(12):1367–72. doi: 10.1038/nbt.1511
- Cox J, Hein MY, Lubner CA, Paron I, Nagaraj N, Mann M. Accurate Proteome-Wide Label-Free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed MaxLFQ. *Mol Cell Proteomics* (2014) 13(9):2513–26. doi: 10.1074/mcp.M113.031591
- Nogueira LG, Santos RH, Fiorelli AI, Mairena EC, Benvenuti LA, Bocchi EA, et al. Myocardial Gene Expression of T-Bet, GATA-3, Ror- γ t, FoxP3, and Hallmark Cytokines in Chronic Chagas Disease Cardiomyopathy: An Essentially Unopposed TH1-Type Response. *Mediators Inflamm* (2014) 2014:914326. doi: 10.1155/2014/914326

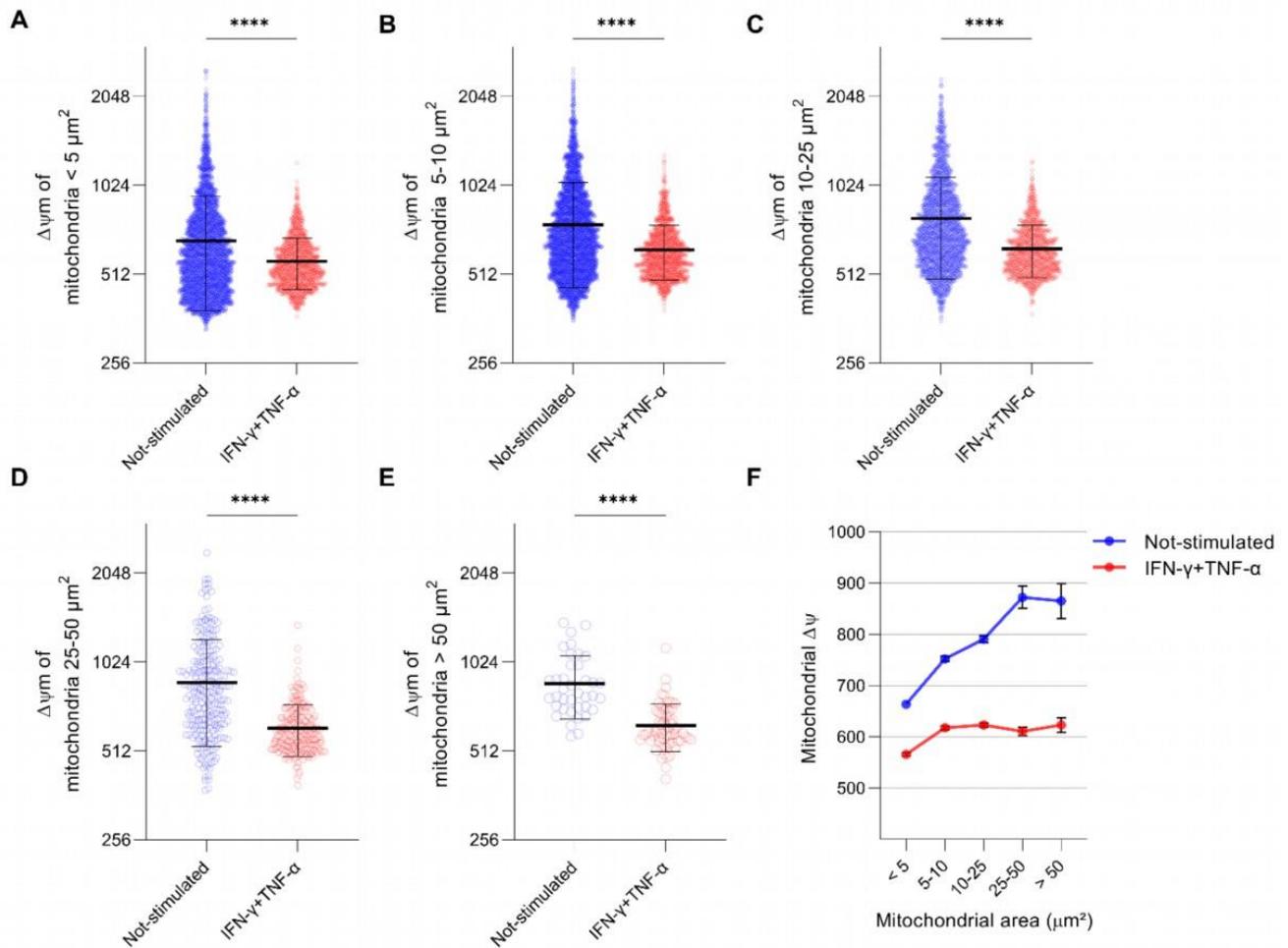
37. Cunha-Neto E, Rizzo LV, Albuquerque F, Abel L, Guilherme L, Bocchi E, et al. Cytokine Production Profile of Heart-Infiltrating T Cells in Chagas' Disease Cardiomyopathy. *Braz J Med Biol Res* (1998) 31(1):133–7. doi: 10.1590/S0100-879X1998000100018
38. Kolwicz SC, Purohit S, Tian R. Cardiac Metabolism and Its Interactions With Contraction, Growth, and Survival of Cardiomyocytes. *Circ Res* (2013) 113(5):603–16. doi: 10.1161/CIRCRESAHA.113.302095
39. Kastl L, Sauer S, Beissbarth T, Becker M, Krammer P, Gülow K. TNF- α Stimulation Enhances ROS-Dependent Cell Migration via NF- κ B Activation in Liver Cells. *Free Radic Biol Med* (2014) 75(Suppl 1):S32. doi: 10.1016/j.freeradbiomed.2014.10.765
40. Doll DN, Rellick SL, Barr TL, Ren X, Simpkins JW. Rapid Mitochondrial Dysfunction Mediates TNF- α -Induced Neurotoxicity. *J Neurochem* (2015) 132(4):443–51. doi: 10.1111/jnc.13008
41. Gottlieb E, Vander Heiden MG, Thompson CB. Bcl-X(L) Prevents the Initial Decrease in Mitochondrial Membrane Potential and Subsequent Reactive Oxygen Species Production During Tumor Necrosis Factor α -Induced Apoptosis. *Mol Cell Biol* (2000) 20(15):5680–9. doi: 10.1128/MCB.20.15.5680-5689.2000
42. Cao YY, Chen ZW, Gao YH, Wang XX, Ma JY, Chang SF, et al. Exenatide Reduces Tumor Necrosis Factor- α -Induced Apoptosis in Cardiomyocytes by Alleviating Mitochondrial Dysfunction. *Chin Med J (Engl)* (2015) 128(23):3211–8. doi: 10.4103/0366-6999.170259
43. Maiti AK, Sharba S, Navabi N, Forsman H, Fernandez HR, Lindén SK. IL-4 Protects the Mitochondria Against Tnf α and Ifn γ Induced Insult During Clearance of Infection With *Citrobacter Rodentium* and *Escherichia Coli*. *Sci Rep* (2015) 5:15434. doi: 10.1038/srep15434
44. Cheshire JL, Baldwin AS. Synergistic Activation of NF- κ B by Tumor Necrosis Factor α and Gamma Interferon via Enhanced I κ B Alpha Degradation and De Novo I κ B Alpha Degradation. *Mol Cell Biol* (1997) 17(11):6746–54. doi: 10.1128/MCB.17.11.6746
45. Kauppinen A, Suuronen T, Ojala J, Kaarniranta K, Salminen A. Antagonistic Crosstalk Between NF- κ B and SIRT1 in the Regulation of Inflammation and Metabolic Disorders. *Cell Signal* (2013) 25(10):1939–48. doi: 10.1016/j.cellsig.2013.06.007
46. Rakshit S, Chandrasekar BS, Saha B, Victor ES, Majumdar S, Nandi D. Interferon-Gamma Induced Cell Death: Regulation and Contributions of Nitric Oxide, Cjun N-Terminal Kinase, Reactive Oxygen Species and Peroxynitrite. *Biochim Biophys Acta* (2014) 1843(11):2645–61. doi: 10.1016/j.bbamcr.2014.06.014
47. Mihm MJ, Coyle CM, Schanbacher BL, Weinstein DM, Bauer JA. Peroxynitrite Induced Nitration and Inactivation of Myofibrillar Creatine Kinase in Experimental Heart Failure. *Cardiovasc Res* (2001) 49(4):798–807. doi: 10.1016/S0008-6363(00)00307-2
48. Stavnichuk R, Shevalye H, Lupachyk S, Obrosova A, Groves JT, Obrosova IG, et al. Peroxynitrite and Protein Nitration in the Pathogenesis of Diabetic Peripheral Neuropathy. *Diabetes Metab Res Rev* (2014) 30(8):669–78. doi: 10.1002/dmrr.2549
49. Wang F, Zhang S, Jeon R, Vuckovic I, Jiang X, Lerman A, et al. Interferon Gamma Induces Reversible Metabolic Reprogramming of M1 Macrophages to Sustain Cell Viability and Pro-Inflammatory Activity. *EBioMedicine* (2018) 30:303–16. doi: 10.1016/j.ebiom.2018.02.009
50. Koo SJ, Chowdhury IH, Szczesny B, Wan X, Garg NJ. Macrophages Promote Oxidative Metabolism To Drive Nitric Oxide Generation in Response to *Trypanosoma Cruzi*. *Infect Immun* (2016) 84(12):3527–41. doi: 10.1128/IAI.00809-16
51. Gupta S, Bhatia V, Wen JJ, Wu Y, Huang MH, Garg NJ. *Trypanosoma Cruzi* Infection Disturbs Mitochondrial Membrane Potential and ROS Production Rate in Cardiomyocytes. *Free Radic Biol Med* (2009) 47(10):1414–21. doi: 10.1016/j.freeradbiomed.2009.08.008
52. Estrada D, Specker G, Martínez A, Dias PP, Hissa B, Andrade LO, et al. Cardiomyocyte Diffusible Redox Mediators Control. *Biochem J* (2018) 475(7):1235–51. doi: 10.1042/BCJ20170698
53. Roberge S, Roussel J, Andersson DC, Meli AC, Vidal B, Blandel F, et al. TNF- α -Mediated Caspase-8 Activation Induces ROS Production and TRPM2 Activation in Adult Ventricular Myocytes. *Cardiovasc Res* (2014) 103(1):90–9. doi: 10.1093/cvr/cvu112
54. Ide T, Tsutsui H, Kinugawa S, Utsumi H, Kang D, Hattori N, et al. Mitochondrial Electron Transport Complex I Is a Potential Source of Oxygen Free Radicals in the Failing Myocardium. *Circ Res* (1999) 85(4):357–63. doi: 10.1161/01.RES.85.4.357
55. Dutta D, Calvani R, Bernabei R, Leeuwenburgh C, Marzetti E. Contribution of Impaired Mitochondrial Autophagy to Cardiac Aging: Mechanisms and Therapeutic Opportunities. *Circ Res* (2012) 110(8):1125–38. doi: 10.1161/CIRCRESAHA.111.246108
56. Elorza AA, Soffia JP. mtDNA Heteroplasmy at the Core of Aging-Associated Heart Failure. An Integrative View of OXPHOS and Mitochondrial Life Cycle in Cardiac Mitochondrial Physiology. *Front Cell Dev Biol* (2021) 9:625020. doi: 10.3389/fcell.2021.625020
57. Yan C, Duanmu X, Zeng L, Liu B, Song Z. Mitochondrial DNA: Distribution, Mutations, and Elimination. *Cells* (2019) 8(4):1–15. doi: 10.3390/cells8040379
58. Lewis W, Day BJ, Kohler JJ, Hosseini SH, Chan SS, Green EC, et al. Decreased mtDNA, Oxidative Stress, Cardiomyopathy, and Death From Transgenic Cardiac Targeted Human Mutant Polymerase Gamma. *Lab Invest* (2007) 87(4):326–35. doi: 10.1038/labinvest.3700523
59. Blasco N, Cámara Y, Núñez E, Beà A, Barés G, Forné C, et al. Cardiomyocyte Hypertrophy Induced by Endonuclease G Deficiency Requires Reactive Oxygen Radicals Accumulation and Is Inhibitable by the Micropeptide Humanin. *Redox Biol* (2018) 16:146–56. doi: 10.1016/j.redox.2018.02.021
60. Cunha-Neto E, Teixeira PC, Fonseca SG, Bilate AM, Kalil J. Myocardial Gene and Protein Expression Profiles After Autoimmune Injury in Chagas' Disease Cardiomyopathy. *Autoimmun Rev* (2011) 10(3):163–5. doi: 10.1016/j.autrev.2010.09.019
61. Garg N, Gerstner A, Bhatia V, DeFord J, Papaconstantinou J. Gene Expression Analysis in Mitochondria From Chagasic Mice: Alterations in Specific Metabolic Pathways. *Biochem J* (2004) 381(Pt 3):743–52. doi: 10.1042/BJ20040356
62. Ouarhache M, Marquet S, Frade AF, Ferreira AM, Ianni B, Almeida RR, et al. Rare Pathogenic Variants in Mitochondrial and Inflammation-Associated Genes May Lead to Inflammatory Cardiomyopathy in Chagas Disease. *J Clin Immunol* (2021) 41(July 2021):1048–63. doi: 10.1007/s10875-021-01000-y
63. Ba X, Gupta S, Davidson M, Garg NJ. *Trypanosoma Cruzi* Induces the Reactive Oxygen Species-PARP-1-RelA Pathway for Up-Regulation of Cytokine Expression in Cardiomyocytes. *J Biol Chem* (2010) 285(15):11596–606. doi: 10.1074/jbc.M109.076984
64. Levick SP, Goldspink PH. Could Interferon-Gamma be a Therapeutic Target for Treating Heart Failure? *Heart Fail Rev* (2014) 19(2):227–36. doi: 10.1007/s10741-013-9393-8
65. Ni C, Ma P, Wang R, Lou X, Liu X, Qin Y, et al. Doxorubicin-Induced Cardiotoxicity Involves Ifn γ -Mediated Metabolic Reprogramming in Cardiomyocytes. *J Pathol* (2019) 247(3):320–32. doi: 10.1002/path.5192
66. Buoncervello M, Maccari S, Ascione B, Gambardella L, Marconi M, Spada M, et al. Inflammatory Cytokines Associated With Cancer Growth Induce Mitochondria and Cytoskeleton Alterations in Cardiomyocytes. *J Cell Physiol* (2019) 234(11):20453–68. doi: 10.1002/jcp.28647
67. Gupta S, Dhiman M, Wen JJ, Garg NJ. ROS Signalling of Inflammatory Cytokines During *Trypanosoma Cruzi* Infection. *Adv Parasitol* (2011) 76:153–70. doi: 10.1016/B978-0-12-385895-5.00007-4
68. Cuadrado A, Manda G, Hassan A, Alcaraz MJ, Barbas C, Daiber A, et al. Transcription Factor NRF2 as a Therapeutic Target for Chronic Diseases: A Systems Medicine Approach. *Pharmacol Rev* (2018) 70(2):348–83. doi: 10.1124/pr.117.014753
69. Sivandzade F, Prasad S, Bhalerao A, Cucullo L. NRF2 and NF- κ B Interplay in Cerebrovascular and Neurodegenerative Disorders: Molecular Mechanisms and Possible Therapeutic Approaches. *Redox Biol* (2019) 21:101059. doi: 10.1016/j.redox.2018.11.017
70. Saha S, Buttari B, Panieri E, Profumo E, Saso L. An Overview of Nrf2 Signaling Pathway and Its Role in Inflammation. *Molecules* (2020) 25(22):1–31. doi: 10.3390/molecules25225474
71. Li X, Liu J, Lu Q, Ren D, Sun X, Rousselet T, et al. AMPK: A Therapeutic Target of Heart Failure-Not Only Metabolism Regulation. *Biosci Rep* (2019) 39(1):1–13. doi: 10.1042/BSR20181767
72. Salminen A, Hyttinen JM, Kaarniranta K. AMP-Activated Protein Kinase Inhibits NF- κ B Signaling and Inflammation: Impact on Healthspan and Lifespan. *J Mol Med (Berl)* (2011) 89(7):667–76. doi: 10.1007/s00109-011-0748-0

73. Sánchez-Villamil JP, Bautista-Niño PK, Serrano NC, Rincon MY, Garg NJ. Potential Role of Antioxidants as Adjunctive Therapy in Chagas Disease. *Oxid Med Cell Longev* (2020) 2020:9081813. doi: 10.1155/2020/9081813
74. Wan X, Wen JJ, Koo SJ, Liang LY, Garg NJ. SIRT1-Pgc1 α -Nfkb Pathway of Oxidative and Inflammatory Stress During Trypanosoma Cruzi Infection: Benefits of SIRT1-Targeted Therapy in Improving Heart Function in Chagas Disease. *PLoS Pathog* (2016) 12(10):e1005954. doi: 10.1371/journal.ppat.1005954
75. Nogueira LG, Santos RH, Ianni BM, Fiorelli AI, Mairena EC, Benvenuti LA, et al. Myocardial Chemokine Expression and Intensity of Myocarditis in Chagas Cardiomyopathy Are Controlled by Polymorphisms in CXCL9 and CXCL10. *PLoS Negl Trop Dis* (2012) 6(10):e1867. doi: 10.1371/journal.pntd.0001867
76. Jones WK, Brown M, Ren X, He S, McGuinness M. NF-kappaB as an Integrator of Diverse Signaling Pathways: The Heart of Myocardial Signaling? *Cardiovasc Toxicol* (2003) 3(3):229–54. doi: 10.1385/CT:3:3:229
77. Karwi QG, Uddin GM, Ho KL, Lopaschuk GD. Loss of Metabolic Flexibility in the Failing Heart. *Front Cardiovasc Med* (2018) 5:68. doi: 10.3389/fcvm.2018.00068
78. Silva JS, Vespa GN, Cardoso MA, Aliberti JC, Cunha FQ. Tumor Necrosis Factor Alpha Mediates Resistance to Trypanosoma Cruzi Infection in Mice by Inducing Nitric Oxide Production in Infected Gamma Interferon-Activated Macrophages. *Infect Immun* (1995) 63(12):4862–7. doi: 10.1128/iai.63.12.4862-4867.1995
79. Vila-del Sol V, Punzón C, Fresno M. IFN-Gamma-Induced TNF-Alpha Expression Is Regulated by Interferon Regulatory Factors 1 and 8 in Mouse Macrophages. *J Immunol* (2008) 181(7):4461–70. doi: 10.4049/jimmunol.181.7.4461
80. Ramos GC, van den Berg A, Nunes-Silva V, Weirather J, Peters L, Burkard M, et al. Myocardial Aging as a T-Cell-Mediated Phenomenon. *Proc Natl Acad Sci USA* (2017) 114(12):E2420–9. doi: 10.1073/pnas.1621047114
81. Houssari M, Dumesnil A, Tardif V, Kivelä R, Pizzinat N, Boukhalfa I, et al. Lymphatic and Immune Cell Cross-Talk Regulates Cardiac Recovery After Experimental Myocardial Infarction. *Arterioscler Thromb Vasc Biol* (2020) 40(7):1722–37. doi: 10.1161/ATVBAHA.120.314370

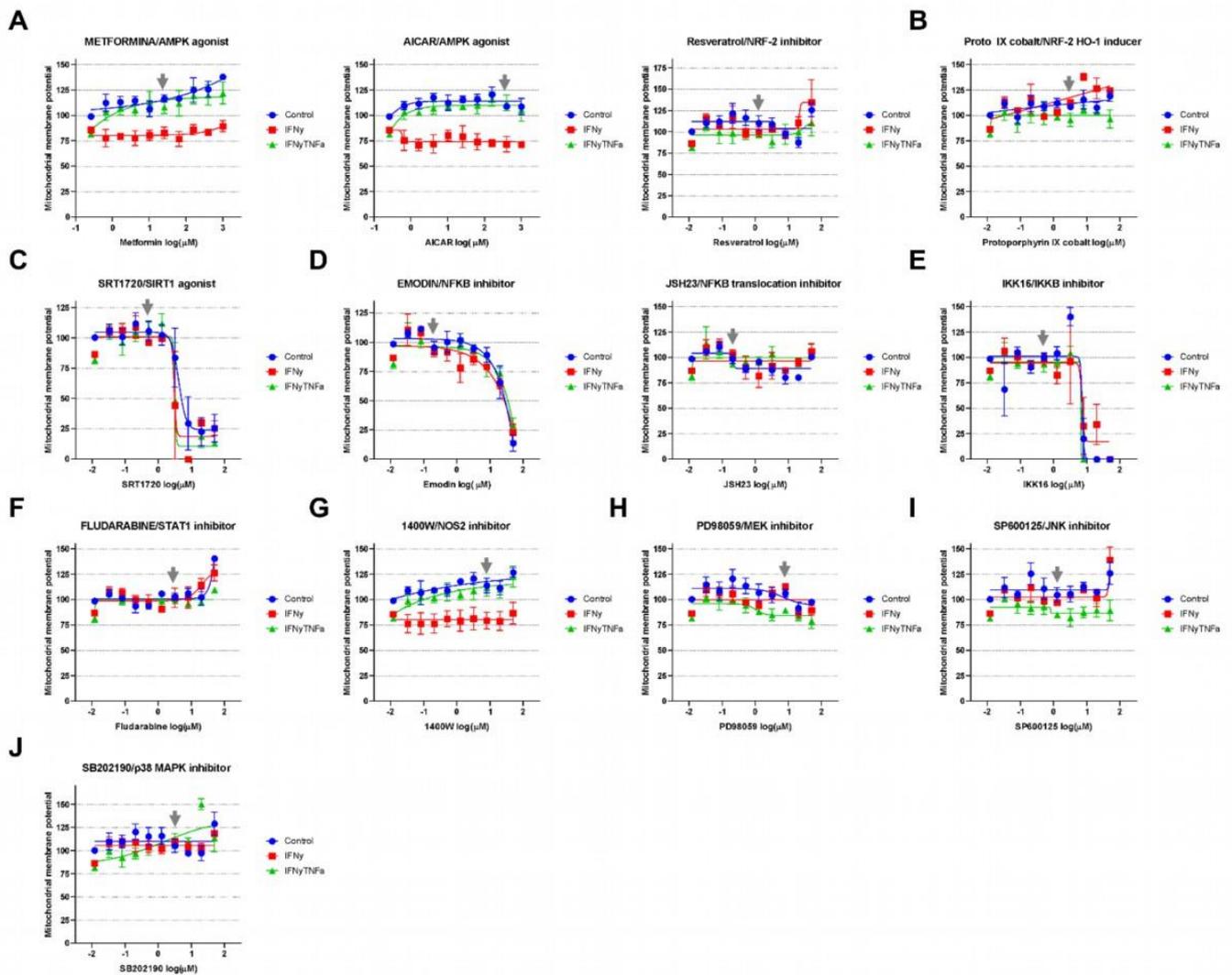
Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

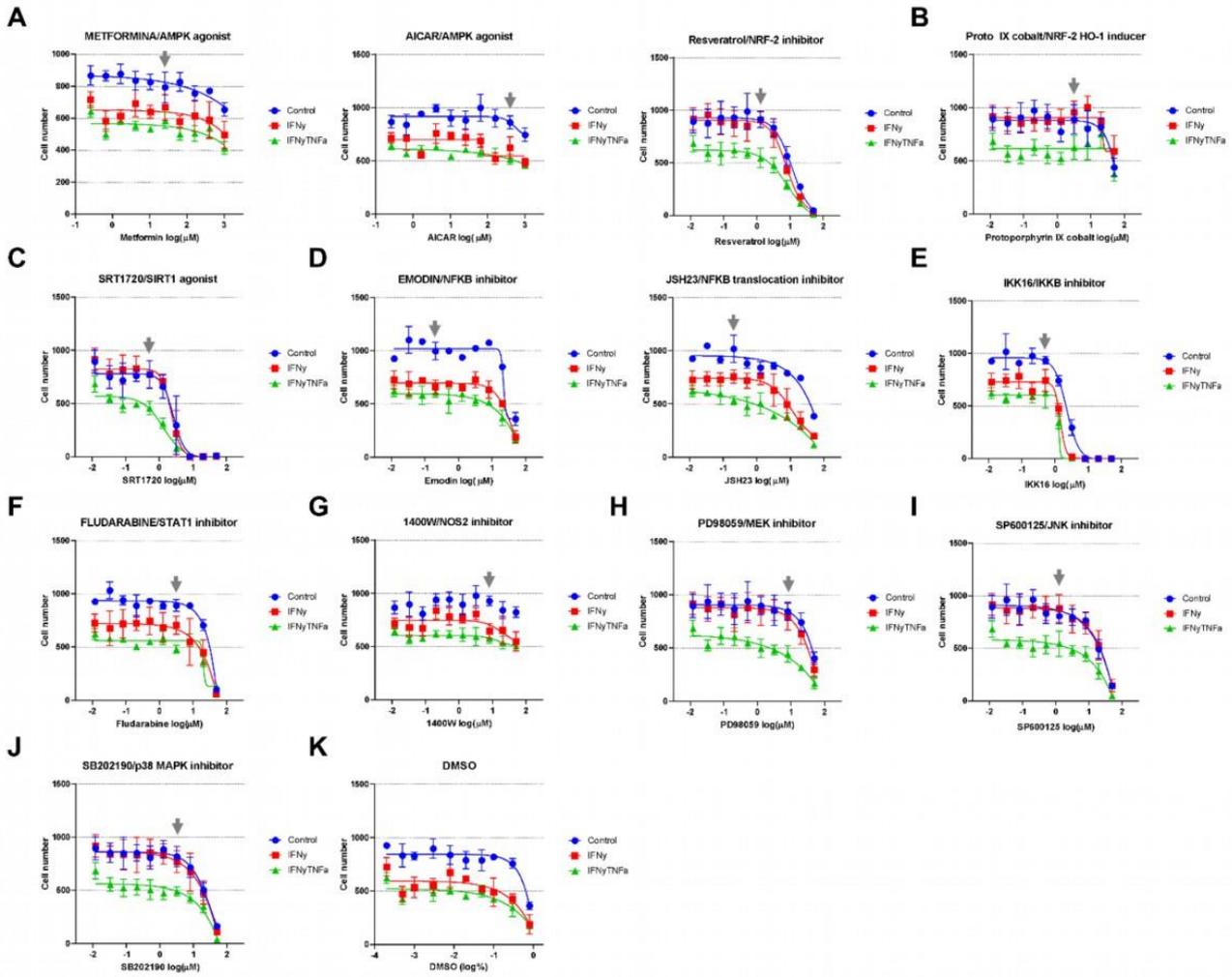
Copyright © 2021 Nunes, Andrieux, Brochet, Almeida, Kitano, Honda, Iwai, Andrade-Silva, Goudenège, Alcântara Silva, Vieira, Levy, Bydlowski, Gallardo, Torres, Bocchi, Mano, Santos, Bacal, Pomerantzeff, Laurindo, Teixeira, Nakaya, Kalil, Procaccio, Chevillard and Cunha-Neto. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



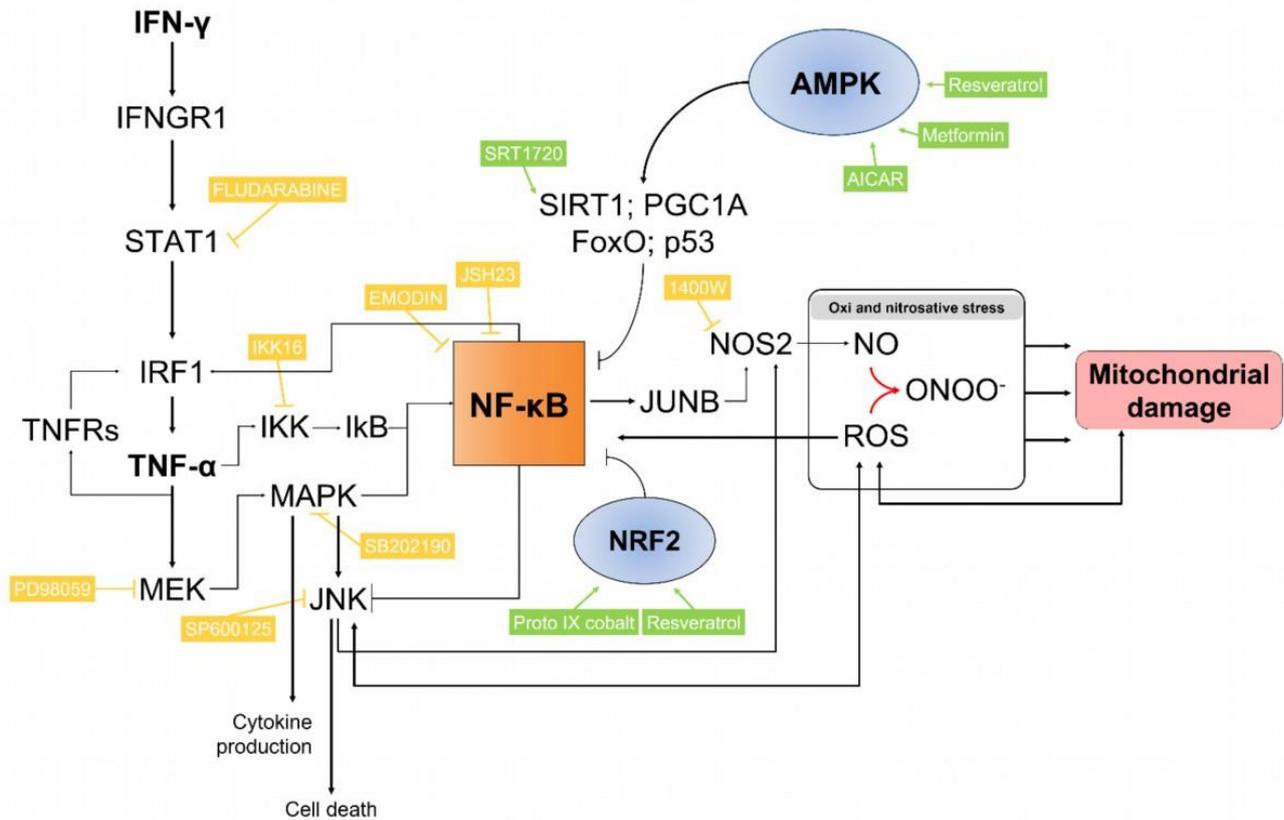
Supplementary Figure 1. Stratification of mitochondria area and membrane potential after stimulation with IFN- γ and TNF- α in human cardiomyocytes AC-16 cells. Cells were stimulated with 10 ng/ml of IFN- γ + 5 ng/ml of TNF- α for 48 hours and then multi-labelled with TMRM, Mitotracker DeepRed, NucGreen and DAPI. TMRM fluorescence was measured when colocalized with the fluorescence of Mitotracker Deep Red in live cells (NucGreen-negative). Mitochondria were stratified based on their area in: lower than 5 (A), between 5-10 (B), between 10-25 (C), between 25-50 (D) and (E) higher than 50 μm^2 . (F) average (\pm SEM) value obtained from the mitochondrial membrane potential of each stratification of mitochondrial size. **** $p < 0.0001$; Mann Whitney test.



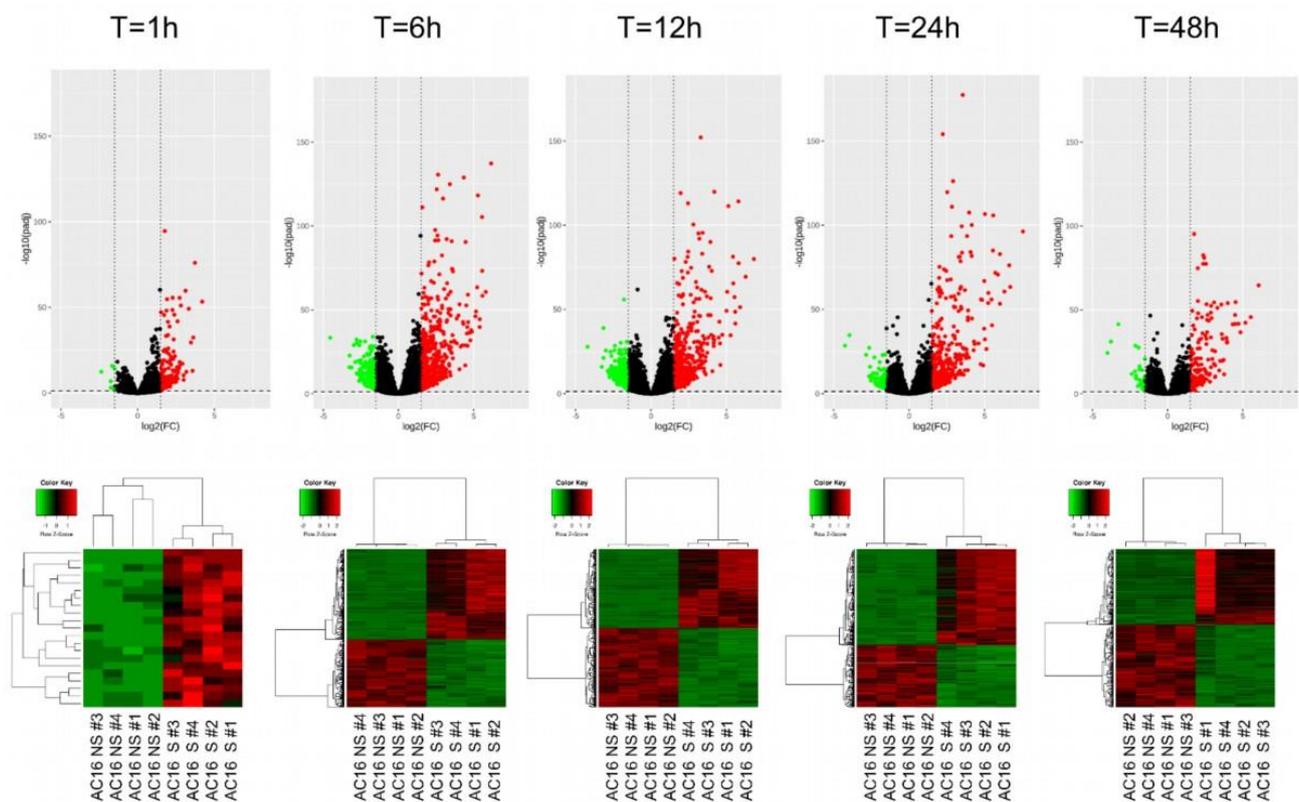
Supplementary Figure 2. Compound screening in AC-16 mitochondrial membrane potential. AC-16 cells were treated with serially diluted agonists of AMPK (A), NRF2 (B), SIRT1 (C) or inhibitors of NF- κ B (D), IKK β (E), STAT1 (F), NOS2 (G), MEK1 and MEK2 (H), JNK (I) and MAPK (J) alone (blue line) or in combination with 10 ng/ml IFN- γ (red line) or 10 ng/ml of IFN- γ plus 5 ng/ml TNF- α (green line) for 48 hours. Specific doses were selected based on the highest effect on mitochondrial $\Delta\Psi_m$ and less than 10% loss on cell number. The first dot of the titration (to the left) is without compounds. Grey arrows indicate concentrations selected. All data are shown as percentage to not-stimulated cells. Standard error of the mean is from 3 independent experiments.



Supplementary Figure 3. Compound screening in the cell number of AC-16. AC-16 cells were treated with serially diluted agonists of AMPK (**A**), NRF2 (**B**), SIRT1 (**C**) or inhibitors of NF- κ B (**D**), IKK β (**E**), STAT1 (**F**), NOS2 (**G**), MEK1 and MEK2 (**H**), JNK (**I**), MAPK (**J**) and DMSO (**K**) alone (blue line) or in combination with 10 ng/ml of IFN- γ (red line) or 10 ng/ml of IFN- γ plus 5 ng/ml of TNF- α (green line) for 48 hours. The first dot of each titration (to the left) is without compounds. Grey arrows indicate concentrations selected. Standard error of the mean is from 3 independent experiments.



Supplementary Figure 4. Schematic representation of the inhibitors and agonists and their targets in the IFN- γ and TNF- α signaling pathways. The pro-inflammatory cytokine IFN- γ interacts with its transmembrane receptor IFNGR1 and signals mainly through the signal transducer and activator of transcription 1 (STAT1) and Interferon Regulatory Factor 1 (IRF1) intracellular transduction pathway to achieve transcriptional activation of IFN- γ -inducible genes, such as TNF- α and Nuclear Factor Kappa B (NF- κ B). Activation of NF- κ B leads to an upregulation of Nitric oxide synthase 2 (NOS2) gene, which in turn increases the amount of intracellular nitric oxide (NO), reacting to reactive oxygen species (ROS) to generate peroxynitrite (ONOO⁻). Arrows indicate activation and flat line means inhibitory interaction. While ONOO⁻ is important for pathogenic response, it causes damages to mitochondria, leading to reduction of mitochondria membrane potential and fragmentation. Proteins such AMPK and NRF2 are involved in the cellular response to oxidative stress countering the damaging effects of NF- κ B. Compounds highlighted in yellow are inhibitors and green are agonists.



Supplementary Figure 5. Transcriptomic analysis on AC-16 cardiomyocyte cell line stimulated with IFN- γ and TNF- α . AC-16 cells were stimulated with IFN- γ and TNF- α during 1h to 48h. Gene expression analysis was done between the various time points taking as reference the T=0h time point. At each time is provided the volcano plot and the heatmap. Each stimulation was performed 4 times.

Toutes les tables supplémentaires sont directement **accessibles en ligne** à l'adresse suivante :

<https://www.frontiersin.org/articles/10.3389/fimmu.2021.755862/full>

Annexe 2

Chagas Disease Megaesophagus Patients Carrying Variant MRPS18B P260A Display Nitro-Oxidative Stress and Mitochondrial Dysfunction in Response to IFN- γ Stimulus

Karla Deysiree Alcântara Silva*, João Paulo Silva Nunes*, Pauline Andrieux, Pauline Brochet, Rafael Ribeiro Almeida, Andréia Cristina Kazue Kuramoto Takara, Natalia Bueno Pereira, Laurent Abel, Aurelie Cobat, Ricardo Costa Fernandes Zaniratto, Débora Levy, Sergio Paulo Bydlowski, Ivan Cecconello, Francisco Carlos Bernal da Costa Seguro, Jorge Kalil, Christophe Chevillard* and Edecio Cunha-Neto*.

Biomedicines. 2022 Sep 7;10(9):2215. doi: 10.3390/biomedicines10092215. PMID: 36140315; PMCID: PMC9496350.

Présentation de la publication :

Les CCC sont les formes chroniques symptomatiques les plus fréquentes de la maladie de Chagas. Toutefois, 10% des patients développent des formes digestives, appelées mégacolon ou mégaoesophage, en fonction de l'organe atteint. Globalement, 30 à 40% des maladies d'origine mitochondriale progressent en cardiomyopathies, et 15% en troubles du système digestif et du système nerveux. Compte tenu des similitudes avec le devenir des patients atteints de la maladie de Chagas, ainsi que de l'importance déjà prouvée de la mitochondrie dans le développement des CCC, nous avons conduit une analyse de séquençage de l'exome (WES) sur des échantillons sanguins de patients atteints des formes digestives de la maladie de Chagas et de patients asymptomatiques.

En se limitant aux gènes impliqués dans les processus mitochondriaux, nous avons réussi à identifier deux variants associés de manière significative aux formes digestives, lesquels sont localisés dans les gènes MRPS18B (rs34315095) et FAM185A (rs201667800). Le gène MRPS18B code pour une protéine liée au ribosome mitochondrial, ce qui en fait un acteur d'intérêt dans le contexte de cette étude. La présence du variant rs34315095 a également été confirmée par le séquençage de Sanger.

Afin de mieux comprendre l'impact du variant rs34315095 sur la pathologie, nous avons stimulé les cellules de patients atteints de formes digestives porteurs de cette mutation (homozygotes mutés et hétérozygotes) ainsi que celles des non mutés (homozygotes non mutés) avec de l'IFN- γ . Les résultats montrent que la production de NO induite après stimulation était significativement plus élevée chez les patients mutés (hétérozygotes et homozygotes). De plus, les patients homozygotes mutés ont montré une augmentation significative de la production de ROS mitochondriaux, et à l'inverse une diminution de la production d'ATP.

L'ensemble de ces résultats met en évidence l'existence d'une composante génétique mitochondriale qui favorise le développement des formes chroniques digestives, soulignant ainsi l'importance cruciale de la mitochondrie dans la maladie de Chagas.

Contribution personnelle :

J'ai participé aux analyses statistiques de cette étude, ainsi qu'à l'écriture aux corrections de l'article, dont je suis 4^{ème} auteur.



Article

Chagas Disease Megaesophagus Patients Carrying Variant *MRPS18B* P260A Display Nitro-Oxidative Stress and Mitochondrial Dysfunction in Response to IFN- γ Stimulus

Karla Deysiree Alcântara Silva ^{1,2,3,†} , João Paulo Silva Nunes ^{1,2,3,4,†} , Pauline Andrieux ⁴, Pauline Brochet ⁴ , Rafael Ribeiro Almeida ^{1,3}, Andréia Cristina Kazue Kuramoto Takara ^{1,3}, Natalia Bueno Pereira ^{1,2}, Laurent Abel ^{5,6,7}, Aurelie Cobat ^{5,6,7}, Ricardo Costa Fernandes Zaniratto ¹ , Débora Levy ¹ , Sergio Paulo Bydlowski ¹ , Ivan Ceconello ⁸, Francisco Carlos Bernal da Costa Seguro ⁸, Jorge Kalil ^{1,2,3}, Christophe Chevillard ^{4,*} and Edecio Cunha-Neto ^{1,2,3,*}



Citation: Silva, K.D.A.; Nunes, J.P.S.; Andrieux, P.; Brochet, P.; Almeida, R.R.; Kuramoto Takara, A.C.K.; Pereira, N.B.; Abel, L.; Cobat, A.; Zaniratto, R.C.F.; et al. Chagas Disease Megaesophagus Patients Carrying Variant *MRPS18B* P260A Display Nitro-Oxidative Stress and Mitochondrial Dysfunction in Response to IFN- γ Stimulus. *Biomedicines* **2022**, *10*, 2215. <https://doi.org/10.3390/biomedicines10092215>

Academic Editor: Marie-Pierre Golinelli-Cohen

Received: 1 August 2022

Accepted: 1 September 2022

Published: 7 September 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

- ¹ Laboratory of Immunology, Heart Institute (Incor) Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, São Paulo 05403-900, Brazil
- ² Division of Clinical Immunology and Allergy, Faculdade de Medicina da Universidade de São Paulo, São Paulo 05403-000, Brazil
- ³ Institute for Investigation in Immunology/INCT, São Paulo 05403-900, Brazil
- ⁴ Institut MarMaRa, INSERM, UMR_1090, Aix Marseille Université, TAGC Theories and Approaches of Genomic Complexity, 13288 Marseille, France
- ⁵ Laboratory of Human Genetics of Infectious Diseases, Necker Branch, INSERM U1163, Necker Hospital for Sick Children, 75015 Paris, France
- ⁶ Imagine Institute, University of Paris, 75015 Paris, France
- ⁷ St. Giles Laboratory of Human Genetics of Infectious Diseases, Rockefeller Branch, The Rockefeller University, New York, NY 10065, USA
- ⁸ Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, São Paulo 05403-900, Brazil
- * Correspondence: christophe.chevillard@univ-amu.fr (C.C.); edecunha@gmail.com (E.C.-N.)
- † These authors contributed equally to this work.

Abstract: Chagas disease (CD), caused by the protozoan parasite *Trypanosoma cruzi*, affects 8 million people, and around 1/3 develop chronic cardiac (CCC) or digestive disease (megaesophagus/megacolon), while the majority remain asymptomatic, in the indeterminate form of Chagas disease (ASY). Most CCC cases in families with multiple Chagas disease patients carry damaging mutations in mitochondrial genes. We searched for exonic mutations associated to chagasic megaesophagus (CME) in genes essential to mitochondrial processes. We performed whole exome sequencing of 13 CME and 45 ASY patients. We found the damaging variant *MRPS18B* 688C > G P230A, in five out of the 13 CME patients (one of them being homozygous; 38.4%), while the variant appeared in one out of 45 ASY patients (2.2%). We analyzed the interferon (IFN)- γ -induced nitro-oxidative stress and mitochondrial function of EBV-transformed lymphoblastoid cell lines. We found the CME carriers of the mutation displayed increased levels of nitrite and nitrated proteins; in addition, the homozygous (G/G) CME patient also showed increased mitochondrial superoxide and reduced levels of ATP production. The results suggest that pathogenic mitochondrial mutations may contribute to cytokine-induced nitro-oxidative stress and mitochondrial dysfunction. We hypothesize that, in mutation carriers, IFN- γ produced in the esophageal myenteric plexus might cause nitro-oxidative stress and mitochondrial dysfunction in neurons, contributing to megaesophagus.

Keywords: chagasic megaesophagus; mitochondria; mitochondrial mutation; mitochondrial dysfunction; interferon-gamma; *MRPS18B*; Chagas disease; *Trypanosoma cruzi*

1. Introduction

Chagas disease is caused by the flagellate protozoan *Trypanosoma cruzi* and the transmission occurs through the bite of triatomine vectors [1]. Patients infected with *T. cruzi*

develop acute and chronic forms of the disease. The acute phase is characterized by high parasitemia and increased production of the pro-inflammatory cytokine interferon-gamma (IFN- γ) [2]. In the chronic phase, a proportion of the infected patients remain asymptomatic, without cardiac or digestive damage—in the so-called indeterminate form of Chagas disease (ASY)—while 30% of patients end up developing serious damage to the heart—chronic Chagas cardiomyopathy (CCC)—associated with heart failure, cardiac electrical conduction disorders and cardiac arrhythmias [3]. About 10% of patients develop digestive tract motility disorders, most often dilatations in the esophagus and colon (megaesophagus and megacolon, respectively). The digestive forms of Chagas disease cause a considerable decrease in quality of life, often requiring surgical intervention [4,5]. In both cases, there is degeneration of the myenteric plexuses and destruction of neurons that control the motility of that segment of the digestive tract [6].

The differential susceptibility of 30% of those infected by *T. cruzi* to evolution to CCC and 10% to digestive disease, while most patients remain in the indeterminate form, suggests the participation of host genetics. This possibility was reinforced by the finding of family aggregation of CCC cases [7]. More than 150 case–control studies on CCC have been published, studying common genetic polymorphisms in over 70 candidate genes, mainly related to the immune response [8]. Only two studies, however, evaluated genetic polymorphisms associated with digestive forms of Chagas disease, ficolin-2 [9] and HLA-G [10].

The scarcity of data highlights the importance of studying the genetic aspects of the digestive form of Chagas disease. Recent results from our group pioneered the use of the whole exome sequencing (WES) methodology to investigate rare genetic variants associated with CCC in families with multiple cases of Chagas disease [11]. We observed that, in each family, patients with CCC shared rare heterozygous, non-synonymous and pathogenic variants among them. Such genetic variants were absent from family members with ASY, without cardiac involvement. Of the 25 CCC specific pathogenic variants observed, 11 occurred in mitochondrial genes and 14 in inflammatory genes; CCC-specific mitochondrial gene variants were found in five out of the six families studied.

Mitochondria play a central role in energy supply, metabolic regulation and cell death in response to harmful stimuli [12]. Mitochondriopathies, caused by homozygous pathogenic variants in genes encoded in the nucleus and mitochondrial DNA, are the most common monogenic syndromes. They severely impair energy metabolism and cause mitochondrial dysfunction and functional deterioration in tissues with high metabolic demand, such as cardiac, skeletal muscle, nervous tissue and liver [13,14]. It is interesting to notice that 30–40% of cases of mitochondrial diseases progress to cardiomyopathy, cardiac conduction disorders and severe arrhythmias [15,16] and 15% for digestive motility disorders, including megacolon and megaesophagus with denervation of the myenteric nervous plexuses, and autonomic nervous system disorders [17,18]. The similarity in the nature and prevalence of cardiac, digestive and autonomic nervous system involvement in mitochondriopathies and those observed in the evolution of Chagas disease suggests that heterozygous pathogenic genetic variants, which cause a partial reduction in mitochondrial function, may play a role in the differential evolution of Chagas disease, both for CCC and for the digestive forms.

Studies show that inflammatory cytokines such as IFN- γ and TNF- α cause mitochondrial damage in several cell types [19,20]. IFN- γ and TNF- α , which are abundantly produced in Chagas disease in response to the perennial stimulus of *T. cruzi* [2,21,22], cause a reduction in oxidative metabolism, in the expression of energy metabolism enzymes, in the activity of lipid beta-oxidation and in the production of mitochondrial ATP, and reduction in the action of mitochondrial membrane potential ($\Delta\Psi_m$) via the NF- κ B-dependent signaling pathway [3,23–29]. IFN- γ induces superoxide generation by increasing the expression of NADPH oxidases [30] and induces NO generation by upregulating nitric oxide synthase (NOS2) [31]. Superoxide reacts with NO to produce the highly reactive nitrogen species peroxynitrite, which can destroy *T. cruzi* but also can cause host cell and

mitochondrial damage. We hypothesized that pathogenic mutations in genes essential to mitochondrial processes were associated with the digestive form of Chagas disease. To evaluate our hypothesis, we performed whole exome sequencing in carriers of CME and the indeterminate form of Chagas disease. To assess whether cells from patients carrying these mutations were more likely to suffer cytokine-induced nitro-oxidative stress and mitochondrial dysfunction, we also evaluated the effect of IFN- γ in the Epstein–Barr virus-immortalized B lymphoblastoid cells (EBV-LCL) obtained from patients with CME carrying or not carrying mutations in mitochondrial genes. These cell lines have been used to study mitochondrial dysfunction in genetic mitochondriopathies [32].

2. Methods

2.1. Patients and Blood Collection

Patients with Chagas disease megaesophagus (CME) from the Division of Digestive Tract Surgery of Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (HC-FMUSP) were invited to participate in the study. An informed consent form was applied if the patient wanted to participate in the study. The study was approved by ethical committee (by the number 35658720.2.0000.0068) and all patients selected for this study had at least two serological tests for Chagas disease (ELISA and/or indirect immunofluorescence—IFI) performed by the clinical analysis laboratory of HC-FMUSP.

The control group was made up of patients with the indeterminate form of Chagas disease (ASY) recruited at the Chagas Disease Outpatient Clinic at the General Cardiomyopathies Center of the Heart Institute (InCor), Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo.

2.2. Whole Exome Sequencing (WES) of Patients with CME and ASY

DNA was extracted from whole blood of CD patients with the QIAamp DNA Blood Maxi Kit (Qiagen, Redwood City, CA, USA). The exome of 13 CME patients and 45 ASY patients was sequenced by Genewiz (Paris, France). Low-quality adapter sequences and nucleotides were removed using Trimmomatic v.0.36. The contigs were aligned to the reference genome *Homo sapiens* GRCh37 using the Edico Genome Dragen's pipeline. Variant call format (VCF) files were used for the annotations and prioritization of mutations in the VarAFT software (version 2.06, Windows 10 platform) (<https://varaft.eu>; 29 April 2019) which generates a file for each individual with a list of all variants detected from several databases such as 1000 Genomes and ExAC. Pathogenic variants were performed using the following criteria based on the markers that were considered significant: maintenance of variants located in exon and splicing zones, elimination of synonymous variants, four different databases (SIFT, UMD and two PolyPhen) and the selected mutation considered pathogenic in at least 3 of them. The MRPS18B gene of CME and ASY patients was amplified by real-time PCR using the following primers: F 5'-TACCTCTCATGGGGCTGTGAGT-3' and R 5'-TTCCGCTGGTGTCTAGGGGG-3'. The resulting amplicon was subjected to Sanger sequencing (Genewiz, Leipzig, Germany). EBV-derived lymphoblastoid cell lines (EBV-LCL) from four patients carrying MRPS18B P260A (one homozygous and the other 3 heterozygous) and three wild type CME patients were generated for downstream analyses.

2.3. Variants Selection

An average of 140,677 annotated variants were detected in each CME patient whereas an average 136,177 annotated variants were detected in each asymptomatic subject. In total, 456,344 variants were detected in at least one of the 13 CME patients whereas 646,398 variants were detected in at least one of the 37 ASY subjects.

According to our hypothesis, we focused our analysis only on genes that are involved in mitochondrial functions. The mitochondrial gene list was a combination of all genes contained in the Mitochondrion Gene Ontology term and Mitocarta 2.0 (see Supplemental Table S1). This gene list contains 1531 genes which can be included either in the nuclear genome or in the specific mitochondrial circular genome. Among the CME

variants, a small fraction ($6.6\% = 30,184/456,344$) of the variants were included in our candidate genes.

Among these candidate variants, we selected the variants located in coding (exonic or splicing) regions and that were non-synonymous. For each of these variants, we looked for pathogenicity data available in four databases (Shift, PolyPhen2_HDIV_pred, PolyPhen2_HVAR_pred and UMD predictor). We selected the variants described as pathogenic (or damaging) in at least three of these databases. So, the number of candidate variants dropped to 217 pathogenic non-synonymous variants located in exonic or splicing regions of one of our candidate genes.

In 2014, the CADD score was introduced as a comprehensive tool that aims to take the results of many known prediction tools into account. Here, we focus our analysis on variants with a CADD score above 20. A scaled CADD score of 20 means that a variant is amongst the top 1% of deleterious variants in the human genome. In our statistical analysis, we compared the number of mutated chromosomes in each group. The cutoff value for statistic tests was fixed at 0.005.

2.4. Isolation of Peripheral Blood Mononuclear Cells

Peripheral blood mononuclear cells (PBMCs) were purified from 20 mL of blood collected in tubes treated with heparin from CME patients. Blood was diluted in 0.9% saline solution (1:1 dilution) and carefully superimposed on Ficoll-Hypaque (Amersham Biosciences, Uppsala, Sweden) and submitted to centrifugation at $800 \times g$ (without braking and without acceleration) for 30 min for the separation of PBMCs from red blood cells and plasma. The PBMCs were collected from the interphase with the aid of a Pasteur pipette and washed 3 times in 0.9% saline at $300 \times g$ 8 min. Cells were counted and subjected to immortalization using Epstein–Barr virus.

2.5. Production of a Supernatant Rich in Epstein–Barr Virus Particles

To enrich the supernatant with Epstein–Barr virus (EBV) particles, approximately 2×10^6 monkey B95.8 cells purchased from Banco de Células do Rio de Janeiro (BCRJ) were grown in specific medium R15 (RPMI 1640 (Gibco®, New York, NY, USA) supplemented with 15% inactivated FBS (Life Technologies™, New York, NY, USA), 10 U/mL penicillin, 100 µg/mL streptomycin and 0.25 µg/mL amphotericin B (Thermo Fisher Scientific, Waltham, MA, USA)) at a density of 0.2×10^6 cells per ml in R15 medium in a flask with 25 cm² of growth surface (T25). The cells were kept in an incubator at 37 °C with 5% CO₂ for seven days and the medium was transferred to a new 75 cm² flask pre-filled with 10 mL of R15 medium. Once a week, 10 mL of R15 was added until reaching 50 mL. Then, the flask was completely closed and kept in the incubator at 37 °C for 14 days. The medium was then collected and filtered through a 0.22 µm filter to eliminate debris. Aliquots of 1 mL of the supernatant were frozen and stored at -80 °C.

2.6. Establishment of EBV-LCL from Patients

The PBMCs isolated from all patients were infected with the EBV for transformation of B cells into EBV-LCL. EBV-LCL are continuous sources of cells for genetic and functional studies [33] and have been widely used for the study of mitochondrial function in mitochondrial diseases [32,34]. Transformation was carried out as follows: 1.0×10^6 viable PBMCs were incubated in 400 µL of R20 medium (RPMI 1640 (Gibco®, New York, NY, USA) supplemented with 20% inactivated FBS (Life Technologies™, New York, NY, USA), 10 U/mL of penicillin, 100 µg/mL of streptomycin and 0.25 µg/mL of amphotericin B (Thermo Fisher Scientific, Waltham, MA, USA), with 200 µL of the EBV-enriched supernatant and 2.4 µL of cyclosporin A 1.0 mg/mL (Sigma-Aldrich, Saint Louis, MI, USA) in a 1.9 cm² surface plate (24-well plate). The cells were incubated at 37 °C with 5% CO₂ for seven days without disturbance.

On the 7th day, 250 µL of R20 medium was added to each well and incubated for another seven days in an incubator at 37 °C with 5% CO₂ without disturbance. On the

14th day, the entire content of the three wells was transferred to a 12.5 cm² flask (T12.5) pre-filled with 1 mL of R20 and the cells were incubated at 37 °C with 5% CO₂. Every two or three days, 200–500 µL of R20 was gradually added depending on the growth rate of the cells.

When the culture volume reached 4 mL, cells were centrifuged at 100× *g* for 10 min at room temperature and then the supernatant was discarded. Cells were suspended at a density of 0.4 × 10⁶ viable cells/mL in R15 medium and left to grow until 20 × 10⁶ cells. Then, cells were cryopreserved with 90% FBS (Gibco®, New York, NY, USA) and 10% DMSO (Sigma-Aldrich, Saint Louis, MI, USA) in liquid nitrogen for further analyses.

2.7. Evaluation of mtROS in LCLs

Twenty-four hours before the experiment, the EBV-LCL cells from each patient were centrifuged at 100× *g* for 10 min and the counting was carried out with a Neubauer chamber and trypan blue 0.4% (Invitrogen, New York, NY, USA). The cells were kept at a concentration of 0.4 × 10⁶ viable cells per ml of R15.

On the day of the experiment, the cells were centrifuged at 100× *g* for 10 min, counted and seeded at a concentration of 15 × 10⁴/mL in 150 µL of R15 containing or not containing interferon-γ (IFN-γ, 25 ng/mL) in sterile 96-well U-bottom plates. After stimulation, cells were transferred to V-bottom 96-well plates, centrifuged at 300× *g* for 5 min and suspended in Hanks saline buffer with Ca²⁺ and Mg²⁺ pH 7.4 (HBSS++) containing 5 µM of MitoSOX™ Red (Thermo Fisher Scientific, Waltham, MA, USA), specific for mitochondrial superoxide. The cells were then incubated in the dark at 37 °C for 10 min.

Afterwards, the plate was centrifuged at 300× *g* for 5 min and the cells were resuspended in 150 µL of HBSS++ containing 3.5 µM of Hoechst 33342 (ThermoFisher Scientific, Waltham, MA, USA) and 1 drop per ml of NucGreen (ReadyProbes, Thermo Fisher Scientific, Waltham, MA, USA). The stained cells were transferred to 96-well black plates (Corning 3603) at 3 × 10⁴ per well in 100 µL of HBSS++ and centrifuged at 300× *g* for 2 min without acceleration and braking. Images of five fields per well were captured in a high-content screening system (HCS, ImageXpress, Molecular Devices, Sunnyvale, CA, USA) at 200× magnification using Hoechst 33342 (nucleus), FITC (dead cell nucleus) and Texas red (mtROS) filters.

The MitoSOX signal was measured using the software Columbus. First, we identified the total number of cells by using the nucleus probe Hoechst 33342. Then, dead cells were excluded by using a second probe that stains the nucleus green in dead cells. By excluding dead cells, we measured MitoSOX fluorescence intensity inside live cells at approximately 580 nm (Texas red). The contrast intensity (*c*) is automatically calculated by Columbus with the following function of the mean intensity in the region (*a*), and the mean intensity in the neighborhood of the region (*b*): $c = (a - b)/(a + b)$. Approximately 9000–10,000 cells were measured per patient.

2.8. Sample Preparation of EBV-LCL Stimulated with IFN-γ

For concurrent quantification of nitrite, ATP and nitrated proteins, 50 × 10⁴ cells from each patient were seeded in 2.5 mL of R15 without phenol red in 6-well plates. Cells were treated with IFN-γ (25 ng/mL) for 48 h. After treatment, cells were collected, divided into two 1.5 mL microtubes and centrifuged at 800× *g* for 5 min. Conditioned supernatant was collected in 1.5 mL microtubes and stored in a freezer at −80 °C until use for nitrite quantification.

The pellets formed in the two 1.5 mL microtubes were processed in different ways on ice: (1) the first cell pellet was lysed in 200 µL of TE buffer (100 mM Tris and 4 EDTA, pH = 7.5). Afterwards, the lysate was heated at 95 °C for 7 min and centrifuged at 20,800× *g* for 3 min and the supernatant was collected and then stored in a freezer at −80 °C until use for ATP quantification; (2) the second cell pellet was lysed in 200 µL of lysis buffer (10 mM HEPES, 0.32 M sucrose, 0.1 mM EDTA, 1.0 mM dithiothreitol (DTT), 5.25 µL leupeptin, 0.3 µM aprotinin and 125 µg/mL PMSF at pH 7.4) with the aid of a cell disruptor

(VirSonic 100 (Virtis, Warminster, PA, USA)) for 3 cycles of 10 s with 15 s intervals on ice and centrifuged at $10,000\times g$ for 30 min at 4 °C and the supernatant was collected for quantification of nitrated proteins.

For both lysates, protein concentration was determined by the Bradford method (Bio-Rad, Hercules, CA, USA) and samples were stored at $-80\text{ }^{\circ}\text{C}$ until use.

2.9. Quantification of Nitrite, ATP and Nitrated Protein

The quantification of nitrite in the supernatant was performed using the Griess Reagent Kit (Thermo Fisher Scientific, Waltham, MA, USA) and ATP production was quantified using the ATP Determination Kit (Thermo Fisher Scientific, Waltham, MA, USA) as per the manufacturer's instructions.

For the quantification of nitrated proteins, 5 μg of the protein was vacuum added to a nitrocellulose membrane 0.45 μm with the aid of a HYBRID DOT MANIFOLD dot blot apparatus (Bethesda Research Laboratories, Bethesda, MD, USA). The membrane was air dried in an Eppendorf ThermoMixer[®] thermoblock (Thermo Fisher Scientific, Waltham, MA, USA) at 60 °C for 1 h and then blocked with 5% Blotting-Grade Blocker (Bio-Rad, Hercules, CA, USA) + TBS-Tween (NaCl 120 mM, Tris 20 mM, Tween20 0.05%) for 2 h at room temperature (RT), with agitation. After, membrane was washed 3 times for 5 min under agitation with TBS-T. The membrane was then incubated with 1:1000 of the primary antibody 3 free nitrotyrosine (3-NT) in BSA 0.3% + TBS-T overnight at 4 °C under agitation. After, membrane was washed three times and incubated with 1:10,000 of secondary antibody IRDye 680RD Donkey anti-mouse 700 nm (LI-COR, Lincoln, NE, USA) in 0.3% BSA + TBS-T for 2 h at RT, under agitation and protected from light. After, membrane was washed three times again and fluorescence was acquired using the Odyssey[®] Infrared Imaging System scanner (LI-COR, Lincoln, NE, USA) in the 700 nm channel. The normalization was carried out using Ponceau staining in the ImageQuant LAS-400 image reader (GE Healthcare, Chicago, IL, USA).

2.10. Statistical Analysis

Data and statistical analysis were performed using the program GraphPad Prism 8 (Windows 10). A Mann–Whitney test of the mean row statistics in the quantification of mtROS and a Mann–Whitney test for the quantification of nitrite, ATP and nitrate proteins for each patient were carried out. After, a second analysis was carried out per group, C/C (wild type), C/G (heterozygous) and G/G (homozygous), and the mean of each analysis was used. For this second analysis, we used an ANOVA two-way test. For exome analysis, two statistical tests were used: Fisher test and chi-squared. Values of $p < 0.05$ were considered significant.

3. Results

3.1. Exome Sequencing Reactions for CME Patients and ASY Subjects

We performed whole exome sequencing of 13 CME patients and 37 ASY subjects to identify if CME patients display an accumulation of pathogenic polymorphisms in comparison with ASY. The workflow is described in Figure 1.

Two variants were more frequent in patients with Chagas disease megasophagus (*MRPS18B* rs34315095C/G and *FAM185A* rs201667800T/G) (Tables 1–3). The mitochondrial ribosomal protein S18B (*MRPS18B*) gene encodes a 28S subunit protein that belongs to the small subunit of the mitochondrial ribosome. We have little functional information on the *FAM185A* gene. A *FAM185A* variant was found in a study (whole mitochondrial genome sequencing, in conjunction with high-throughput genotyping arrays) carried out to discover genetic variants associated with exercise responses. The link with mitochondria is in fact not obvious [35]. So, we decided to focus on the *MRPS18B* gene, as many homozygous mitochondrial ribosome mutations are associated with genetic mitochondrialopathies.

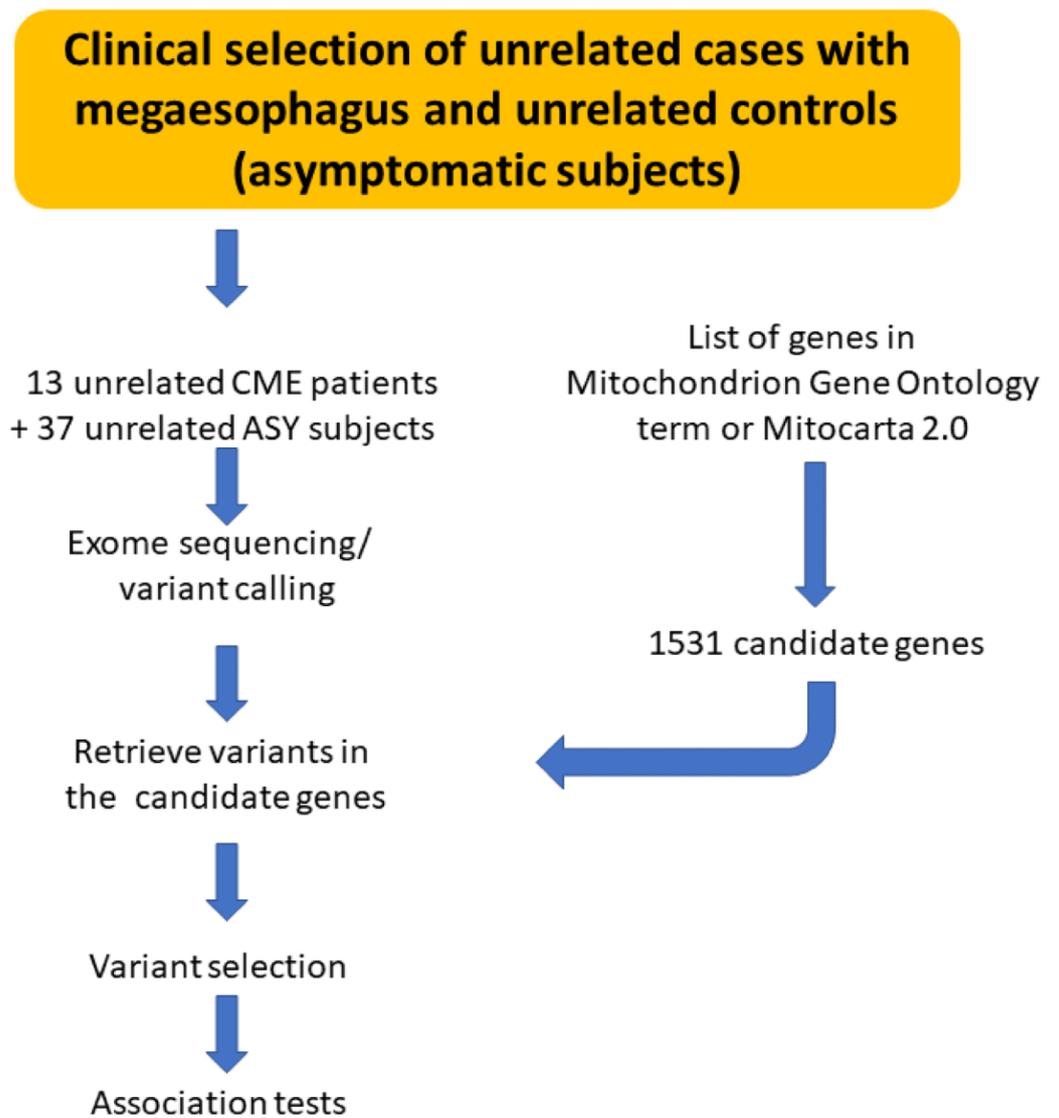


Figure 1. Schematic drawing describing the workflow of analysis.

Table 1. List of associated variants.

| Gene | Ensembl Reference | Chromosome | Position | Ref Allele | Mutated Allele | rs Number | Nucleotide Change (Transcript) | Amino Acid Change (Transcript) | CADD Score |
|----------------|-------------------|------------|-----------|------------|----------------|-------------|---------------------------------|--------------------------------|------------|
| <i>MRPS18B</i> | ENSG00000204568 | Chr6 | 30593485 | C | G | rs34315095 | ENST00000259873: exon7 688C > G | ENST00000259873: P230A | 27.9 |
| <i>FAM185A</i> | ENSG00000222011 | Chr7 | 102417753 | T | G | rs201667800 | ENST00000409231: exon5 538T > G | ENST00000409231: Y180D | 26.0 |

Table 2. Variants’ distribution between cases and controls.

| Mutations | CME Patients | | | ASY Subjects | | |
|--------------------------------|-----------------------|---------------------|----------------------|-----------------------|---------------------|----------------------|
| | Heterozygote Patients | Homozygote Patients | Non-Mutated Patients | Heterozygote Subjects | Homozygote Subjects | Non-Mutated Subjects |
| <i>MRPS18B</i> rs34315095 C/G | 4 | 1 | 8 | 1 | 0 | 36 |
| <i>FAM185A</i> rs201667800 T/G | 6 | 0 | 7 | 2 | 0 | 35 |

Table 3. Statistical test on mutated chromosomes.

| Mutations | CME Patients | | ASY Subjects | | Fisher Exact Test <i>p</i> -Value |
|--------------------------------|---------------------|-------------------------|---------------------|-------------------------|--------------------------------------|
| | Mutated Chromosomes | Non Mutated Chromosomes | Mutated Chromosomes | Non-Mutated Chromosomes | |
| <i>MRPS18B</i> rs34315095 C/G | 6 | 20 | 1 | 73 | 0.001105404 |
| <i>FAM185A</i> rs201667800 T/G | 6 | 20 | 2 | 72 | 0.003611684 |

3.2. Pathogenic Variant of *MRPS18B* Is Accumulated in CME Patients

The *MRPS18B* variant (688C > G, P230A) was present in 38% of CME patients (five of the 13 CME patients) and 3% of ASY patients (1 of 37) (Table 2). One of the five CME patients carrying the mutation was homozygous (G/G) while the other four CME carriers and the ASY carrier were heterozygous (C/G). So, the G allele frequency in CME is 23.1%, whereas this allelic frequency reaches 1.4% in ASY subjects.

The G variant rs34315095 has a frequency of 4.6% in the 1000 Genomes database. Based on the ALFA allele frequency project, in reference populations the alternative allele (G) frequencies were 1.8% (European reference population), 5.5% (African reference population); 5.2% (African American reference population) and 3.8% (Asian reference population). Recently, the ABraOM repository was built, containing genomic variants obtained with whole exome and whole genome sequencing from SABE, a census-based sample of elderly individuals from São Paulo, Brazil [36]. In both databases (exome sequencing database: 609 individuals and whole genome sequencing: 1171 individuals), the G frequency reaches 4.7%.

The confirmatory sequencing was performed by Sanger sequencing (Genewiz). The DNA was amplified by real-time PCR using specific primers for *MRPS18B*. The resulting amplicon was subjected to Sanger sequencing and results confirmed the presence of the polymorphism 680C > G of the *MRPS18B* gene (Figure 2 shows representative chromatograms).

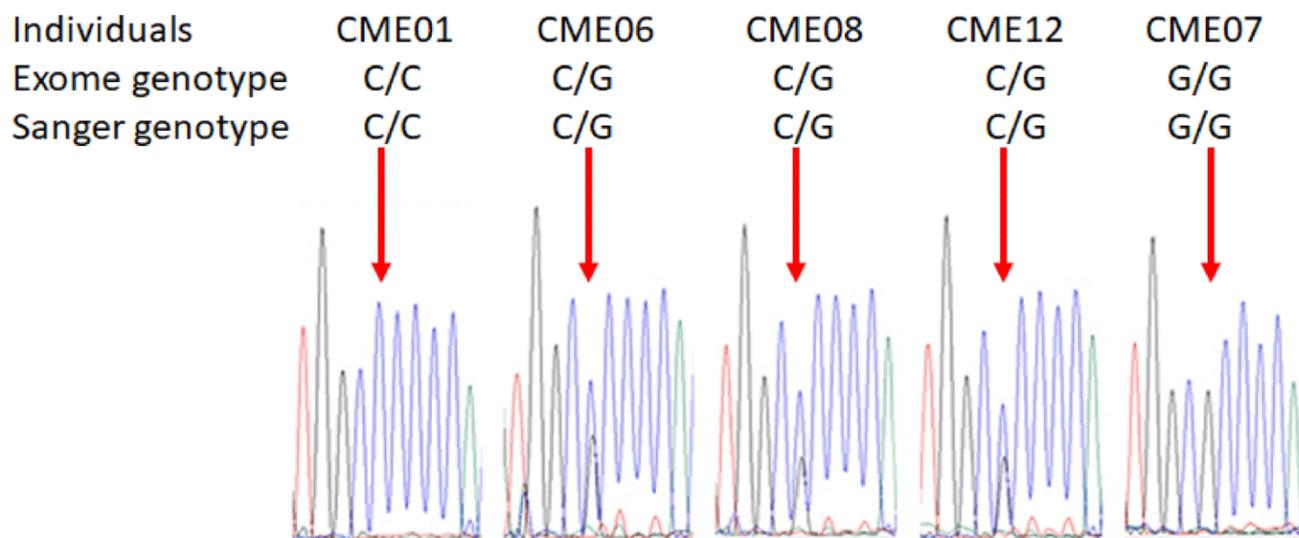


Figure 2. Sanger chromatograms of patients with or without the damaging variant *MRPS18B*. The confirmatory sequencing was performed by Sanger sequencing. The DNA of each patient was extracted using QIAamp DNA Blood Maxi Kit (Qiagen, Germantown, WI, USA) and 25 ng of DNA was used for the amplification of *MRPS18B* gene. PCR quality was assessed by gel electrophoresis and all samples generated a single band corresponding to *MRPS18B*. Then, samples were sequenced by Sanger sequencing by Genewiz. Representative images of Sanger chromatograms of the 5 patients reveal the change in the nucleotide. The red arrow indicates the base change from cytosine to guanine.

Chromatograms of all CME patients are depicted in Supplemental Figure S1.

3.3. IFN- γ Causes Nitro-Oxidative Stress and Mitochondrial Damage in Patients Carrying MRPS18B P260A Variation

In order to investigate if patients carrying the variant 688C > G MRPS18B are more susceptible to IFN- γ 's damaging effects compared to patients carrying the wild type genotype (C/C), we conducted functional assays in EBV-LCL derived from CME heterozygous (C/G) and homozygous (G/G) patients. We quantified mtROS, nitrite, nitrate proteins and synthesis of ATP in non-stimulated and IFN- γ -stimulated cells (25 ng/mL).

We found that C/G (CME06, CME08 and CME12) and G/G (CME07) patients are more susceptible to IFN- γ 's damaging effects compared to C/C (CME01, CME02 and CME03) patients. We identified that C/G and G/G patients had increased nitrite production (31% and 38%, respectively) as compared with the C/C group (Figure 3). Similarly, the C/G patients had a 67% increase and the G/G CME7 patients a 70% increase in nitrated proteins when compared with the C/C group (Figure 3B).

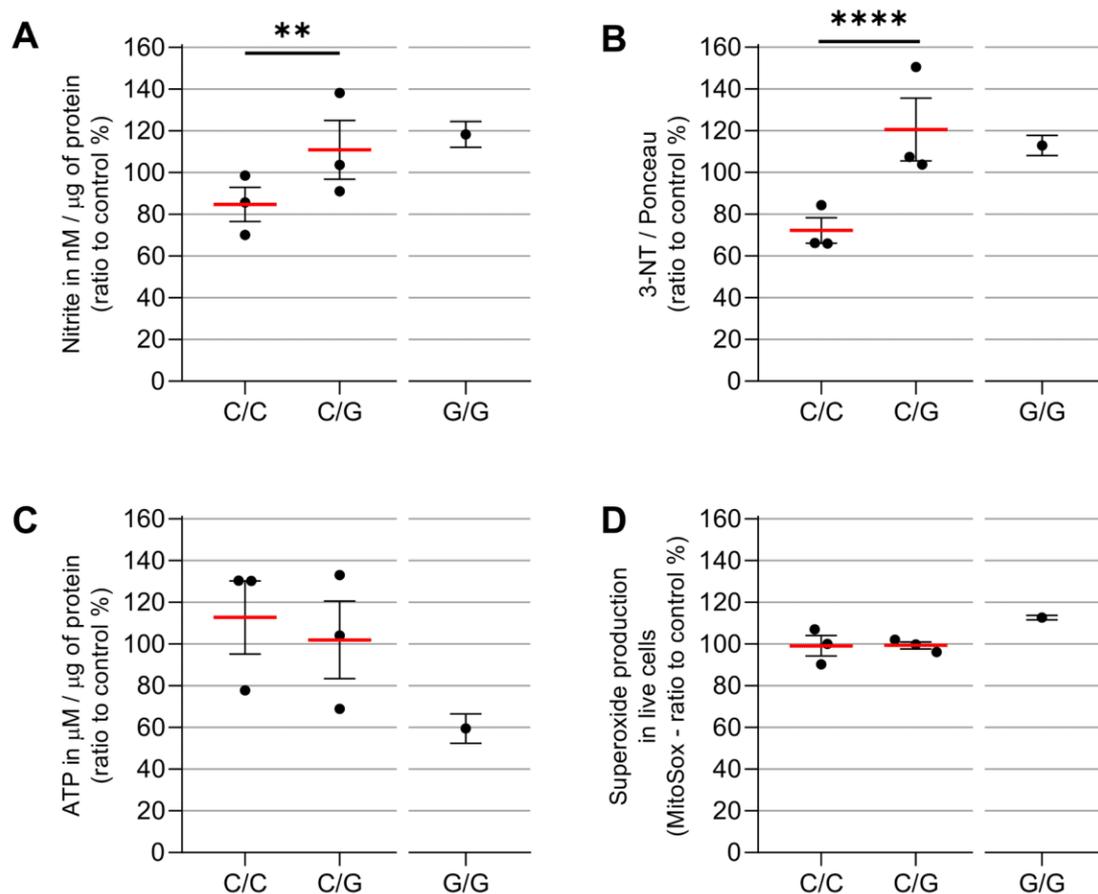


Figure 3. ATP production and nitro-oxidative stress after IFN- γ treatment in carriers or non-carriers of MRPS18B c.688 C > G MRPS18B p. P230A mitochondrial gene mutation. EBV-LCL were obtained from the PBMCs of the patients and infected with EBV for transformation. The EBV-LCL from the patients were stimulated in vitro with IFN- γ . We measured the production of nitrite (A), carried out relative quantification of nitrated proteins (B), ATP (C) and mtROS (D), after the stimulation with IFN- γ 25 ng/mL with MitoSOXTM Red, Griess Reagent Kit, Dot Blot, ATP Determination Kit and the anti-3-nitrotyrosine antibody. We here depict the average ratio between IFN- γ -treated vs. unstimulated cells. Each dot is a patient with ≥ 6 independent measurements. Experiments were repeated several times independently. As we had only one CME patient G/G, statistics were carried out only between C/C patients and C/G patients. Two-way ANOVA statistics were performed between C/C and C/G. ** = $p < 0.01$, **** = $p < 0.0001$.

We next measured ATP production by luciferase-based assay and we found that IFN- γ reduced the ATP production 3% in C/G patients compared to C/C patients, while the CME07 G/G showed a reduction of 54% (Figure 3C) as compared to baseline conditions.

We evaluated mtROS production using MitoSOXTM Red, which binds preferentially to the superoxide radical (O_2^-). Results show there is no significant increase in C/G vs. C/C patients, but homozygous patient CME7 (G/G) showed a 12% increase when compared with the other patients (Figures 3D and 4A). The p value is 0.0022 with a Mann-Whitney test of six independent measurements of IFN-gamma-stimulated cells compared to unstimulated cells. The cell viability assay indicated no cytotoxic effect of IFN- γ at the concentration of 25 ng/mL (Figure 5).

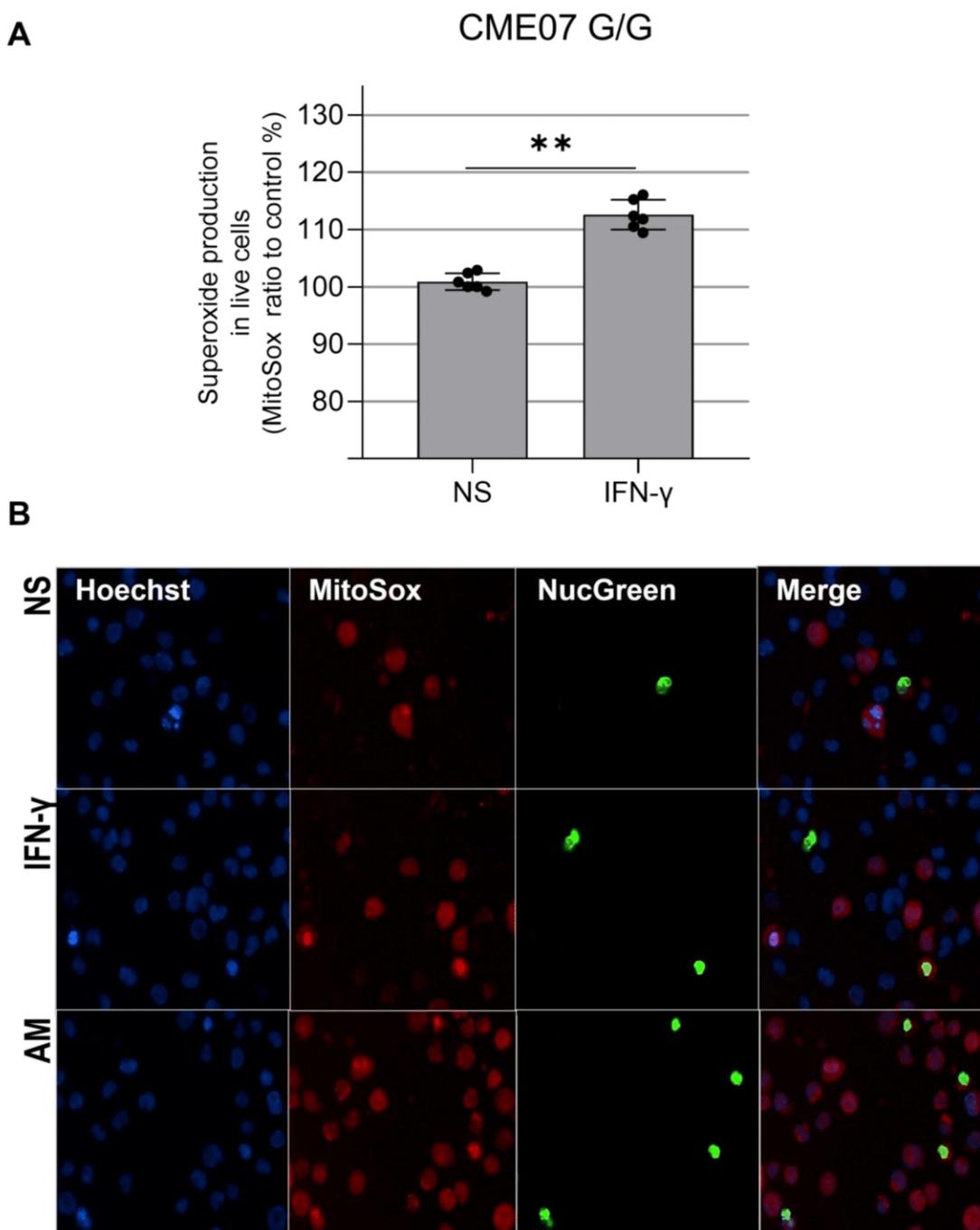


Figure 4. (A). Mitochondrial ROS increase in IFN- γ -treated EBV-LCL from CME07 (homozygous MRPS18B G/G). A total of 15×10^4 EBV-LCL cells were stimulated with IFN- γ 25 ng/mL for 24 h

and cell nuclei were stained with Hoechst 33342 (blue), the nuclei of dead cells were stained with ReadyProbes™ NucGreen (green), also used to measure cell viability, and superoxide was measured with MitoSOX™ (red) (B). Positive control: 560 μ M of antimycin A. Magnification 200 \times . NS: not stimulated. Statistical analysis performed using the Mann–Whitney test. ** = $p < 0.01$.

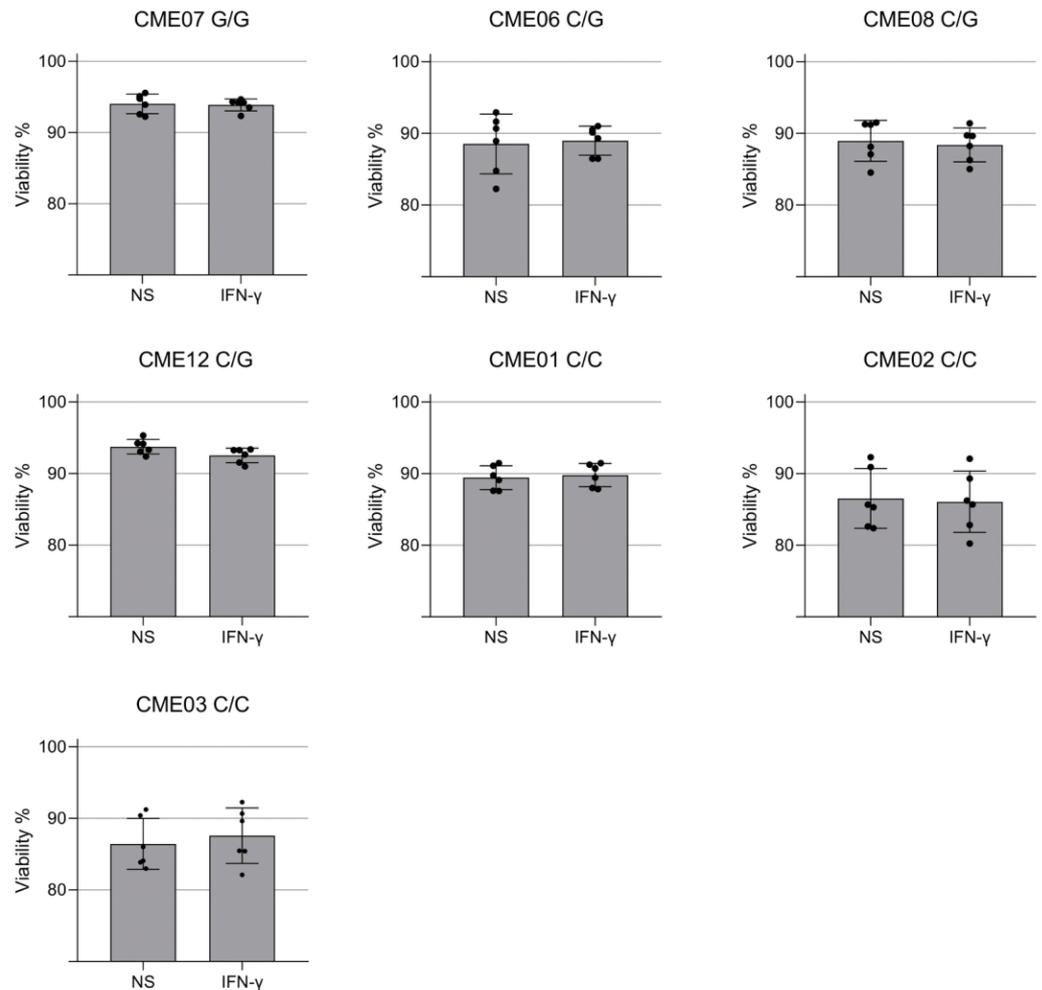


Figure 5. Viability of EBV-LCL cells after treatment with IFN- γ for 24 h. The cell viability of patient-specific EBV-LCL was quantified using the ReadyProbes™ NucGreen dye after 24 h of treatment with IFN- γ (25 ng/mL). Cell viability of EBV-LCL from patients with EMC with the mutation in HOM (G/G), HET (C/G) and non-carriers (C/C). Each dot is an independent measurement. Statistical analysis performed using the Mann–Whitney test.

4. Discussion

This study demonstrated for the first time the relevance of pathogenic variants in patients with Chagas disease megaesophagus. We demonstrate that CME patients have a higher frequency of pathogenic mutations compared to FI patients, such as the mitochondrial mutation MRPS18B 688C > G which was identified in 38% of CME patients and 2% of FI patients. The MRPS18B variant (688C > G, P230A) was characterized by a Sift score of 0.01 (damaging), a PolyPhen score of 0.996 (damaging) and a CADD score of 28. The ENST00000259873: exon7: 688C base is well conserved in the evolution (score 8/9). The proline 230 is located in an external alpha helix.

IFN- γ treatment in EBV-LCL cell lines carrying MRPS18B 688C > G was associated with increased nitro-oxidative stress compared to patients with megaesophagus without

the mutation and the homozygous patient displayed higher damaging effects as shown by decreased ATP production and superoxide production.

In this work, we focused on evaluating the impact of variants of MRPS18B, because recent studies link mitochondrial dysfunction and poor prognosis of Chagas disease. The MRPS18 proteins are encoded by nuclear genes and are involved in protein synthesis within mitochondria. The homologous proteins are zinc-binding proteins that make up the 28S small subunit (MRPS18B and C) and the 39S large subunit (MRPS18A) of the 55S mammalian mitoribosomes [37–40]. Mitochondria have their own translation system for the synthesis of thirteen proteins that build the oxidative phosphorylation complex (OXPHOS). Mutations in mitoribosomal proteins are frequent causes of metabolic and energy disturbances due to deficiencies in mitochondrial protein synthesis [41]. Mutations in the MRPS genes have been shown to have deleterious effects, especially in disorders of the mitochondrial respiratory chain. Mitochondrial disease-causing mutations were found in thirteen small ribosomal subunit mitoribosomal proteins (MRPS1, MRPS2, MRPS6, MRPS7, MRPS9, MRPS11, MRPS14, MRPS16; MRPS22, MRPS23, MRPS25, MRPS34, MRPS39) and four large ribosomal mitochondrial proteins (MRPL3, MRPL12, MRPL24, MRPL44) [41]. Several of them are associated with cardiomyopathy, as was also observed in a study by our group where the presence of a different mutation in the MRPS18B gene was identified in CCC patients [11,41]. Some proteins in this group play extraribosomal roles, such as participating in the regulation of apoptosis and a series of cell signaling cascades [38]. Mutations in the mitochondrial genes that affect mitochondrial translation, such as in mitochondrial tRNAs, have been associated with Leigh syndrome and mitochondrial neurogastrointestinal encephalopathy (MNGIE), two mitochondrial diseases with loss of myenteric plexus neurons and dysphagia [13,18,41,42].

Our results corroborate the role of mitochondrial gene variants in susceptibility to symptomatic forms of Chagas disease, as observed for the first time in CCC patients [11]. However, most mitochondrial diseases affect young children and newborns and often involve multiple organ systems. Interestingly, patient CME07, who carries the homozygous variant, is a 60-year-old patient with digestive and cardiac symptoms (heart failure, New York Heart Association (NYHA) functional class II). The fact that the mitochondrial phenotype in vitro only emerged after cytokine stimulation and was only clinically expressed in an adult is suggestive of a two-hit phenomenon, where the mutation increases the cells' susceptibility to nitro-oxidative stress induced by IFN- γ and mitochondrial dysfunction. Thus, the two-hit mitochondrial damage would be located only in inflammatory sites in patients with Chagas disease; for example, in myenteric plexuses in patients with Chagas disease with megaesophagus [6] and in the heart in the case of CCC [43]. The MRPS18B gene is one of the variant genes identified only in the CCC siblings of one Chagas multicase family, mentioned above [11].

We postulate that CME patients display pathology because this mutation causes mitochondrial ribosome dysfunction, which affects protein synthesis essential for adequate oxidative phosphorylation and mitochondrial homeostasis. Defective oxidative phosphorylation may first lead to increases in ROS which may in turn activate NF- κ B and NOS2, inducing increased levels of NO which under high ROS conditions can generate the powerful oxidant peroxynitrite, causing oxidative damage to the cell and mitochondria.

Chronic immune response is also a key factor associated with worse progression of patients with Chagas disease [44]. The pro-inflammatory cytokines IFN- γ and TNF- α are increased in patients with CCC and digestive disease compared to patients with dilated cardiomyopathy [21,45] and these stimuli are reported to be mitochondrial dysfunction inducers by causing damage to cardiac tissue [3,24,27] and, in in vitro cardiomyocytes [46], promoting an increase in the production of reactive oxygen species (ROS) and nitrogen (RNS) [45,47–49]. Considering that the trigger shared by patients with megaesophagus and Chagas cardiomyopathy is infection with *T. cruzi*, our hypothesis is that these CME patients may also have mitochondrial/energy dysfunction associated with pathology plus mutation in the mitochondrial gene.

We identified that CME patients when exposed to IFN- γ respond differently in the production of cytoplasmic mtROS. Although we did not find differences in the production of mtROS in the group of patients heterozygous for MRPS18B, we found that the CME07 MRPS18B G/G had a significant increase of 12% in mtROS compared to all groups, showing that the homozygous mutation may contribute to a worsening in mitochondrial metabolism that culminates in increased production of mtROS.

A study carried out to evaluate the role of nitric oxide in neuronal destruction in C57BL/6 control mice and knockout for NOS2 and IFN during *T. cruzi* infection demonstrated that the myoenteric denervation observed in the acute phase of protozoan infection is due to the production of nitric oxide (NO) and IFN- γ , as it was observed that control mice produced more NO when compared to the two knockout groups [50]. Our study showed that the EBV-LCL of CME patients carrying the genetic variant in homozygosis and heterozygosis showed an increase in the production of nitrite (product of endogenous NO metabolism) after treatment with IFN- γ , thus showing that the cytokine has an essential role in the increase in nitrite production and, consequently, generating a possible mitochondrial dysfunction and the increase in agents that in large quantities become toxic for the functioning of the mitochondria.

In addition to the increase in nitrite, patients with MRPS18B 688C > G P260A increased the amount of nitrated proteins after stimulation with IFN- γ compared to non-carriers. The identification of nitrated proteins is a consequence of the increase in peroxynitrite (ONOO⁻), a product originated by the combination of superoxide and NO or by the combination of hydrogen peroxide (H₂O₂) with nitrite [51]. Excessive production of ONOO⁻ can lead to tissue damage and mitochondrial dysfunction [51,52]. A study carried out in hepatoma and fibrosarcoma strains also showed that after treatment with IFN- γ there was an increase in peroxynitrite [51].

In previous studies, our group demonstrated that patients with CCC display a reduction in mitochondrial enzymes involved in ATP production [53]. Our investigations showed that this decrease may also be present in CME patients in sites of increased IFN- γ production, where we observed a decrease in ATP synthesis in the EBV-LCL of patients with mitochondrial mutation after treatment with IFN- γ . In fact, the homozygous patient (CME07 G/G) followed this pattern, where IFN- γ stimulation increased mtROS, oxide nitric and tyrosine nitration accompanied by reduced ATP production. The heterozygous group (C/G) increased nitrite and nitrotyrosine in an mtROS-independent way. This could be explained by direct activation of NOS2 by IFN- γ itself or fast catalytic conversion of leaked mtROS into H₂O₂ by antioxidant enzymes such as superoxide dismutase [54]. The increased protein nitration observed in our model could also be explained by combinatory damaging effects between enhanced NF- κ B/NOS2 activation by IFN- γ . The employment of refined methodologies for mtROS detection, such as LC-MS and a “probe-free” electron paramagnetic resonance (EPR)/spin-trapping technique [55], could give us a better clue on mtROS production in these cells.

Our study had limitations. We focused our study on the genes which are essential for mitochondria. We cannot exclude the involvement of additional variants. The small number of analyzed CME patients may mask the identification of additional mitochondrial mutations and the reduced number of LCL used for functional assays presents only suggestive and not definitive results. However, the differences identified in the frequency of genetic variants between the CME (34 variants) and FI groups were robust, and a cellular/molecular phenotype was associated with the variant. To estimate if this pathway could indeed be influencing the EBV-LCL responses of patients, additional tests need to be carried out, such as silencing the pathways associated with genes in the strains and measuring again how the strains would respond to the IFN- γ stimulus.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/biomedicines10092215/s1>, Figure S1: Sanger sequencing chromatograms on CME patients. Table S1: List of the genes contained in the Mitochondrion Gene Ontology term and Mitocarta 2.0.

Author Contributions: Conceptualization, C.C. and E.C.-N.; methodology, K.D.A.S., J.P.S.N.; R.R.A., N.B.P., A.C.K.K.T., F.C.B.d.C.S., P.A., P.B., R.C.F.Z., D.L., L.A. and A.C. writing—original draft preparation, K.D.A.S. and J.P.S.N.; writing—review and editing, E.C.-N., C.C., S.P.B., I.C. and J.K.; funding acquisition, C.C. and E.C.-N. All authors have read and agreed to the published version of the manuscript.

Funding: This work received specific support from the funding body Fundação de Amparo à Pesquisa do Estado de S. Paulo (FAPESP 2019/22858-3). This work was supported by the Institut National de la Santé et de la Recherche Médicale (INSERM); Aix-Marseille University (grant number: AMIDEX “International_2018” MITOMUTCHAGAS); the French Agency for Research (Agence Nationale de la Recherche—ANR) (grant numbers: “Br-Fr-Chagas”, “landscardio”); the Brazilian Council for Scientific and Technological Development (CNPq); the National Institutes of Health/USA (grant numbers: 2 P50 AI098461-02 and 2U19AI098461-06). This work was funded by the Inserm Cross-Cutting Project GOLD. This project has received funding from the Excellence Initiative of Aix-Marseille University—A*Midex a French “Investissements d’Avenir programme”—Institute MarMaRa AMX-19-IET-007. J.P.S.N. was a recipient of a MarMaRa fellowship. The funders did not play any role in the study design, data collection and analysis, decision to publish or preparation of the manuscript.

Institutional Review Board Statement: Written informed consent was obtained from all subjects involved in the study.

Informed Consent Statement: Ethical approval for the study was obtained from the Ethics Committee of the Heart Institute, Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (approval number 35658720.2.0000.0068). Informed written consent was obtained from the participants.

Data Availability Statement: All data are available in the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Thakare, R.; Dasgupta, A.; Chopra, S. An update on benznidazole for the treatment of patients with Chagas disease. *Drugs Today* **2018**, *54*, 15–23. [[CrossRef](#)] [[PubMed](#)]
2. Cunha-Neto, E.; Chevillard, C. Chagas disease cardiomyopathy: Immunopathology and genetics. *Mediat. Inflamm.* **2014**, *2014*, 683230. [[CrossRef](#)]
3. Bocchi, E.A.; Bestetti, R.B.; Scanavacca, M.I.; Cunha Neto, E.; Issa, V.S. Chronic Chagas Heart Disease Management: From Etiology to Cardiomyopathy Treatment. *J. Am. Coll. Cardiol.* **2017**, *70*, 1510–1524. [[CrossRef](#)]
4. Teixeira, A.R.; Nascimento, R.J.; Sturm, N.R. Evolution and pathology in chagas disease—a review. *Mem. Inst. Oswaldo Cruz* **2006**, *101*, 463–491. [[CrossRef](#)]
5. Teixeira, A.R.; Nitz, N.; Guimaro, M.C.; Gomes, C.; Santos-Buch, C.A. Chagas disease. *Postgrad. Med. J.* **2006**, *82*, 788–798. [[CrossRef](#)] [[PubMed](#)]
6. Matsuda, N.M.; Miller, S.M.; Evora, P.R. The chronic gastrointestinal manifestations of Chagas disease. *Clinics* **2009**, *64*, 1219–1224. [[CrossRef](#)]
7. Zicker, F.; Smith, P.G.; Netto, J.C.; Oliveira, R.M.; Zicker, E.M. Physical activity, opportunity for reinfection, and sibling history of heart disease as risk factors for Chagas' cardiopathy. *Am. J. Trop. Med. Hyg.* **1990**, *43*, 498–505. [[CrossRef](#)]
8. Chevillard, C.; Nunes, J.P.S.; Frade, A.F.; Almeida, R.R.; Pandey, R.P.; Nascimento, M.S.; Kalil, J.; Cunha-Neto, E. Disease Tolerance and Pathogen Resistance Genes May Underlie Trypanosoma cruzi Persistence and Differential Progression to Chagas Disease Cardiomyopathy. *Front. Immunol.* **2018**, *9*, 2791. [[CrossRef](#)] [[PubMed](#)]
9. Luz, P.R.; Boldt, A.B.; Grisbach, C.; Kun, J.F.; Velavan, T.P.; Messias-Reason, I.J. Association of L-ficolin levels and FCN2 genotypes with chronic Chagas disease. *PLoS ONE* **2013**, *8*, e60237. [[CrossRef](#)]
10. Dias, F.C.; Mendes-Junior, C.T.; Silva, M.C.; Tristao, F.S.; Dellalibera-Joviliano, R.; Moreau, P.; Soares, E.G.; Menezes, J.G.; Schmidt, A.; Dantas, R.O.; et al. Human leucocyte antigen-G (HLA-G) and its murine functional homolog Qa2 in the Trypanosoma cruzi Infection. *Mediat. Inflamm.* **2015**, *2015*, 595829. [[CrossRef](#)]
11. Ouarhache, M.; Marquet, S.; Frade, A.F.; Ferreira, A.M.; Ianni, B.; Almeida, R.R.; Nunes, J.P.S.; Ferreira, L.R.P.; Rigaud, V.O.; Candido, D.; et al. Rare Pathogenic Variants in Mitochondrial and Inflammation-Associated Genes May Lead to Inflammatory Cardiomyopathy in Chagas Disease. *J. Clin. Immunol.* **2021**, *41*, 1048–1063. [[CrossRef](#)]
12. Rossignol, R.; Malgat, M.; Mazat, J.P.; Letellier, T. Threshold effect and tissue specificity. Implication for mitochondrial cytopathies. *J. Biol. Chem.* **1999**, *274*, 33426–33432. [[CrossRef](#)]
13. Finsterer, J. Mitochondriopathies. *Eur. J. Neurol.* **2004**, *11*, 163–186. [[CrossRef](#)] [[PubMed](#)]

14. Kohda, M.; Tokuzawa, Y.; Kishita, Y.; Nyuzuki, H.; Moriyama, Y.; Mizuno, Y.; Hirata, T.; Yatsuka, Y.; Yamashita-Sugahara, Y.; Nakachi, Y.; et al. A Comprehensive Genomic Analysis Reveals the Genetic Landscape of Mitochondrial Respiratory Chain Complex Deficiencies. *PLoS Genet.* **2016**, *12*, e1005679. [[CrossRef](#)] [[PubMed](#)]
15. El-Hattab, A.W.; Scaglia, F. Mitochondrial Cardiomyopathies. *Front. Cardiovasc. Med.* **2016**, *3*, 25. [[CrossRef](#)] [[PubMed](#)]
16. Scaglia, F.; Towbin, J.A.; Craigen, W.J.; Belmont, J.W.; Smith, E.O.; Neish, S.R.; Ware, S.M.; Hunter, J.V.; Fernbach, S.D.; Vladutiu, G.D.; et al. Clinical spectrum, morbidity, and mortality in 113 pediatric patients with mitochondrial disease. *Pediatrics* **2004**, *114*, 925–931. [[CrossRef](#)]
17. Chelimsky, G.; Shanske, S.; Hirano, M.; Zinn, A.B.; Cohen, M.; McNeeley, K.; Chelimsky, T.C. Achalasia as the harbinger of a novel mitochondrial disorder in childhood. *J. Pediatr. Gastroenterol. Nutr.* **2005**, *40*, 512–517. [[CrossRef](#)]
18. Finsterer, J.; Frank, M. Gastrointestinal manifestations of mitochondrial disorders: A systematic review. *Therap. Adv. Gastroenterol.* **2017**, *10*, 142–154. [[CrossRef](#)]
19. Chen, X.H.; Zhao, Y.P.; Xue, M.; Ji, C.B.; Gao, C.L.; Zhu, J.G.; Qin, D.N.; Kou, C.Z.; Qin, X.H.; Tong, M.L.; et al. TNF-alpha induces mitochondrial dysfunction in 3T3-L1 adipocytes. *Mol. Cell. Endocrinol.* **2010**, *328*, 63–69. [[CrossRef](#)]
20. Lewis, J.A.; Huq, A.; Najarro, P. Inhibition of mitochondrial function by interferon. *J. Biol. Chem.* **1996**, *271*, 13184–13190. [[CrossRef](#)] [[PubMed](#)]
21. Ribeiro, B.M.; Crema, E.; Rodrigues, V., Jr. Analysis of the cellular immune response in patients with the digestive and indeterminate forms of Chagas' disease. *Hum. Immunol.* **2008**, *69*, 484–489. [[CrossRef](#)]
22. Sousa, G.R.; Gomes, J.A.; Fares, R.C.; Damasio, M.P.; Chaves, A.T.; Ferreira, K.S.; Nunes, M.C.; Medeiros, N.I.; Valente, V.A.; Correa-Oliveira, R.; et al. Plasma cytokine expression is associated with cardiac morbidity in chagas disease. *PLoS ONE* **2014**, *9*, e87082. [[CrossRef](#)]
23. Kalovidouris, A.E.; Plotkin, Z.; Graesser, D. Interferon-gamma inhibits proliferation, differentiation, and creatine kinase activity of cultured human muscle cells. II. A possible role in myositis. *J. Rheumatol.* **1993**, *20*, 1718–1723.
24. Lee, H.J.; Oh, Y.K.; Rhee, M.; Lim, J.Y.; Hwang, J.Y.; Park, Y.S.; Kwon, Y.; Choi, K.H.; Jo, I.; Park, S.I.; et al. The role of STAT1/IRF-1 on synergistic ROS production and loss of mitochondrial transmembrane potential during hepatic cell death induced by LPS/d-GalN. *J. Mol. Biol.* **2007**, *369*, 967–984. [[CrossRef](#)]
25. Luss, H.; Watkins, S.C.; Freeswick, P.D.; Imro, A.K.; Nussler, A.K.; Billiar, T.R.; Simmons, R.L.; del Nido, P.J.; McGowan, F.X., Jr. Characterization of inducible nitric oxide synthase expression in endotoxemic rat cardiac myocytes in vivo and following cytokine exposure in vitro. *J. Mol. Cell. Cardiol.* **1995**, *27*, 2015–2029. [[CrossRef](#)]
26. Mariappan, N.; Elks, C.M.; Sriramula, S.; Guggilam, A.; Liu, Z.; Borkhsenius, O.; Francis, J. NF-kappaB-induced oxidative stress contributes to mitochondrial and cardiac dysfunction in type II diabetes. *Cardiovasc. Res.* **2010**, *85*, 473–483. [[CrossRef](#)] [[PubMed](#)]
27. Meyer, A.; Laverny, G.; Allenbach, Y.; Grelet, E.; Ueberschlag, V.; Echaniz-Laguna, A.; Lannes, B.; Alsaleh, G.; Charles, A.L.; Singh, F.; et al. IFN-beta-induced reactive oxygen species and mitochondrial damage contribute to muscle impairment and inflammation maintenance in dermatomyositis. *Acta Neuropathol.* **2017**, *134*, 655–666. [[CrossRef](#)]
28. Ni, C.; Ma, P.; Wang, R.; Lou, X.; Liu, X.; Qin, Y.; Xue, R.; Blasig, I.; Erben, U.; Qin, Z. Doxorubicin-induced cardiotoxicity involves IFN-gamma-mediated metabolic reprogramming in cardiomyocytes. *J. Pathol.* **2019**, *247*, 320–332. [[CrossRef](#)]
29. Wang, D.; McMillin, J.B.; Bick, R.; Buja, L.M. Response of the neonatal rat cardiomyocyte in culture to energy depletion: Effects of cytokines, nitric oxide, and heat shock proteins. *Lab. Invest.* **1996**, *75*, 809–818. [[PubMed](#)]
30. Rada, B.; Leto, T.L. Oxidative innate immune defenses by Nox/Duox family NADPH oxidases. *Contrib. Microbiol.* **2008**, *15*, 164–187. [[CrossRef](#)]
31. Geng, Y.; Hansson, G.K.; Holme, E. Interferon-gamma and tumor necrosis factor synergize to induce nitric oxide production and inhibit mitochondrial respiration in vascular smooth muscle cells. *Circ. Res.* **1992**, *71*, 1268–1276. [[CrossRef](#)] [[PubMed](#)]
32. Van Bergen, N.J.; Crowston, J.G.; Kearns, L.S.; Staffieri, S.E.; Hewitt, A.W.; Cohn, A.C.; Mackey, D.A.; Trounce, I.A. Mitochondrial oxidative phosphorylation compensation may preserve vision in patients with OPA1-linked autosomal dominant optic atrophy. *PLoS ONE* **2011**, *6*, e21347. [[CrossRef](#)]
33. Hussain, T.; Mulherkar, R. Lymphoblastoid Cell lines: A Continuous in Vitro Source of Cells to Study Carcinogen Sensitivity and DNA Repair. *Int. J. Mol. Cell. Med.* **2012**, *1*, 75–87.
34. Castagna, A.E.; Addis, J.; McInnes, R.R.; Clarke, J.T.; Ashby, P.; Blaser, S.; Robinson, B.H. Late onset Leigh syndrome and ataxia due to a T to C mutation at bp 9,185 of mitochondrial DNA. *Am. J. Med. Genet. A* **2007**, *143A*, 808–816. [[CrossRef](#)]
35. Harvey, N.R.; Voisin, S.; Lea, R.A.; Yan, X.; Benton, M.C.; Papadimitriou, I.D.; Jacques, M.; Haupt, L.M.; Ashton, K.J.; Eynon, N.; et al. Investigating the influence of mtDNA and nuclear encoded mitochondrial variants on high intensity interval training outcomes. *Sci. Rep.* **2020**, *10*, 11089. [[CrossRef](#)]
36. Naslavsky, M.S.; Yamamoto, G.L.; de Almeida, T.F.; Ezquina, S.A.M.; Sunaga, D.Y.; Pho, N.; Bozoklian, D.; Sandberg, T.O.M.; Brito, L.A.; Lazar, M.; et al. Exomic variants of an elderly cohort of Brazilians in the ABraOM database. *Hum. Mutat.* **2017**, *38*, 751–763. [[CrossRef](#)]
37. Cavdar Koc, E.; Burkhardt, W.; Blackburn, K.; Moseley, A.; Spremulli, L.L. The small subunit of the mammalian mitochondrial ribosome. Identification of the full complement of ribosomal proteins present. *J. Biol. Chem.* **2001**, *276*, 19363–19374. [[CrossRef](#)] [[PubMed](#)]
38. Gopisetty, G.; Thangarajan, R. Mammalian mitochondrial ribosomal small subunit (MRPS) genes: A putative role in human disease. *Gene* **2016**, *589*, 27–35. [[CrossRef](#)] [[PubMed](#)]

39. Greber, B.J.; Bieri, P.; Leibundgut, M.; Leitner, A.; Aebersold, R.; Boehringer, D.; Ban, N. The complete structure of the 55S mammalian mitochondrial ribosome. *Science* **2015**, *348*, 303–308. [[CrossRef](#)] [[PubMed](#)]
40. Mushtaq, M.; Ali, R.H.; Kashuba, V.; Klein, G.; Kashuba, E. S18 family of mitochondrial ribosomal proteins: Evolutionary history and Gly132 polymorphism in colon carcinoma. *Oncotarget* **2016**, *7*, 55649–55662. [[CrossRef](#)]
41. Ferrari, A.; Del’Olio, S.; Barrientos, A. The Diseased Mitoribosome. *FEBS Lett.* **2021**, *595*, 1025–1061. [[CrossRef](#)]
42. Gorman, G.S.; Chinnery, P.F.; DiMauro, S.; Hirano, M.; Koga, Y.; McFarland, R.; Suomalainen, A.; Thorburn, D.R.; Zeviani, M.; Turnbull, D.M. Mitochondrial diseases. *Nat. Rev. Dis. Primers* **2016**, *2*, 16080. [[CrossRef](#)]
43. Nogueira, L.G.; Santos, R.H.; Fiorelli, A.I.; Mairena, E.C.; Benvenuti, L.A.; Bocchi, E.A.; Stolf, N.A.; Kalil, J.; Cunha-Neto, E. Myocardial gene expression of T-bet, GATA-3, Ror-gammat, FoxP3, and hallmark cytokines in chronic Chagas disease cardiomyopathy: An essentially unopposed TH1-type response. *Mediators Inflamm.* **2014**, *2014*, 914326. [[CrossRef](#)]
44. Reis, D.D.; Jones, E.M.; Tostes, S., Jr.; Lopes, E.R.; Gazzinelli, G.; Colley, D.G.; McCurley, T.L. Characterization of inflammatory infiltrates in chronic chagasic myocardial lesions: Presence of tumor necrosis factor-alpha+ cells and dominance of granzyme A+, CD8+ lymphocytes. *Am. J. Trop. Med. Hyg.* **1993**, *48*, 637–644. [[CrossRef](#)] [[PubMed](#)]
45. Abel, L.C.; Rizzo, L.V.; Ianni, B.; Albuquerque, F.; Bacal, F.; Carrara, D.; Bocchi, E.A.; Teixeira, H.C.; Mady, C.; Kalil, J.; et al. Chronic Chagas’ disease cardiomyopathy patients display an increased IFN-gamma response to *Trypanosoma cruzi* infection. *J. Autoimmun.* **2001**, *17*, 99–107. [[CrossRef](#)] [[PubMed](#)]
46. Nunes, J.P.S.; Andrieux, P.; Brochet, P.; Almeida, R.R.; Kitano, E.; Honda, A.K.; Iwai, L.K.; Andrade-Silva, D.; Goudenege, D.; Alcantara Silva, K.D.; et al. Co-Exposure of Cardiomyocytes to IFN-gamma and TNF-alpha Induces Mitochondrial Dysfunction and Nitro-Oxidative Stress: Implications for the Pathogenesis of Chronic Chagas Disease Cardiomyopathy. *Front. Immunol.* **2021**, *12*, 755862. [[CrossRef](#)]
47. Gupta, S.; Bhatia, V.; Wen, J.J.; Wu, Y.; Huang, M.H.; Garg, N.J. *Trypanosoma cruzi* infection disturbs mitochondrial membrane potential and ROS production rate in cardiomyocytes. *Free Radic. Biol. Med.* **2009**, *47*, 1414–1421. [[CrossRef](#)] [[PubMed](#)]
48. Pissetti, C.W.; Correia, D.; de Oliveira, R.F.; Llaguno, M.M.; Balarin, M.A.; Silva-Grecco, R.L.; Rodrigues, V., Jr. Genetic and functional role of TNF-alpha in the development *Trypanosoma cruzi* infection. *PLoS Negl. Trop. Dis.* **2011**, *5*, e976. [[CrossRef](#)] [[PubMed](#)]
49. Wan, X.; Gupta, S.; Zago, M.P.; Davidson, M.M.; Dousset, P.; Amoroso, A.; Garg, N.J. Defects of mtDNA replication impaired mitochondrial biogenesis during *Trypanosoma cruzi* infection in human cardiomyocytes and chagasic patients: The role of Nrf1/2 and antioxidant response. *J. Am. Heart Assoc.* **2012**, *1*, e003855. [[CrossRef](#)]
50. Arantes, R.M.; Marche, H.H.; Bahia, M.T.; Cunha, F.Q.; Rossi, M.A.; Silva, J.S. Interferon-gamma-induced nitric oxide causes intrinsic intestinal denervation in *Trypanosoma cruzi*-infected mice. *Am. J. Pathol.* **2004**, *164*, 1361–1368. [[CrossRef](#)]
51. Rakshit, S.; Chandrasekar, B.S.; Saha, B.; Victor, E.S.; Majumdar, S.; Nandi, D. Interferon-gamma induced cell death: Regulation and contributions of nitric oxide, cJun N-terminal kinase, reactive oxygen species and peroxynitrite. *Biochim. Biophys. Acta* **2014**, *1843*, 2645–2661. [[CrossRef](#)] [[PubMed](#)]
52. Kauppinen, A.; Suuronen, T.; Ojala, J.; Kaarniranta, K.; Salminen, A. Antagonistic crosstalk between NF-kappaB and SIRT1 in the regulation of inflammation and metabolic disorders. *Cell. Signal.* **2013**, *25*, 1939–1948. [[CrossRef](#)] [[PubMed](#)]
53. Teixeira, P.C.; Santos, R.H.; Fiorelli, A.I.; Bilate, A.M.; Benvenuti, L.A.; Stolf, N.A.; Kalil, J.; Cunha-Neto, E. Selective decrease of components of the creatine kinase system and ATP synthase complex in chronic Chagas disease cardiomyopathy. *PLoS Negl. Trop. Dis.* **2011**, *5*, e1205. [[CrossRef](#)] [[PubMed](#)]
54. MacMillan-Crow, L.A.; Crow, J.P. Does more MnSOD mean more hydrogen peroxide? *Anti-Cancer Agents Med. Chem.* **2011**, *11*, 178–180. [[CrossRef](#)]
55. Cheng, G.; Zielonka, M.; Dranka, B.; Kumar, S.N.; Myers, C.R.; Bennett, B.; Garces, A.M.; Dias Duarte Machado, L.G.; Thiebaut, D.; Ouari, O.; et al. Detection of mitochondria-generated reactive oxygen species in cells using multiple probes and methods: Potentials, pitfalls, and the future. *J. Biol. Chem.* **2018**, *293*, 10363–10380. [[CrossRef](#)] [[PubMed](#)]